Chemoenzymatic Synthesis of Chiral Amines by Carrier Protein-Dependent Enzymes

by

Stephanie W. Chun

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemistry)
in the University of Michigan
2020

Doctoral Committee:

Assistant Professor Alison R. H. Narayan, Chair
Professor Kristina Håkansson
Professor David H. Sherman
Professor Janet L. Smith

Stephanie W. Chun

chusteph@umich.edu

ORCID iD: 0000-0001-9124-9087

To my family, friends and coworkers who supported me on this journey.

Acknowledgements

This work would not have been possible without the support of many, many other kind and amazing who have generously donated their time and assistance to me over several years, but especially in the past five. First, I would like to thank my advisor, Professor Alison Narayan, for her support and guidance. Alison has provided mentorship and new opportunities to me, but also space to pursue and explore some of my more fanciful ideas. My first research position at the University of Michigan was in the lab of Dr. David H. Sherman, where Dr. Matthew DeMars and Dr. Andrew Lowell introduced me to the protein classes that I continued to study for my dissertation. I am grateful to Dr. Sherman for opening his lab to me that summer, and to the rest of my doctoral committee members, Dr. Janet L. Smith and Dr. Kristina Håkansson for their advice and suggestions throughout my graduate career.

I also owe a great deal to those I met at Brandeis University during my undergraduate days. Dr. Isaac Krauss was my undergraduate advisor for three years, training me in the organic chemistry that allowed me to do a good amount of the synthesis in this thesis. My coworkers were also great friends, supporting me on some long nights in lab and sharing wonderful memories together outside of the lab. Our next-door neighbor, Dr. Barry Snider, never failed to brighten the mood and suggested that I apply to the University of Michigan all those years ago (coincidentally, his alma mater).

Starting in a brand new lab is a daunting task, but I am grateful that I was able to experience it with my fellow first-generation lab members, Summer Baker Dockrey, April Lukwoski and Tyler Doyon. My adventure buddy, Attabey Rodríguez Benítez, the other saxitoxin team members,

Sarah Ackenhusen, Dr. Meagan Hinze, Brian Carlson, Duncan Ellinwood and Ye Wang, and the rest of my lab members made it a joy to come to work every day. I survived the first year with the help of my chem bio cluster friends, Daniel Holland-Moritz, Kevon Stanford and Will Walker/

The analyses in this work could not have been completed without the assistance of Dr. Wendy Feng at the Life Sciences Institute and Jim Windak in the Department of Chemistry. I would also like to thank my collaborators and members of other labs who have helped us immensely, Dr. Meredith Skiba, Dr. Gregory Dodge and Yongtong Shero Lao (Smith lab) and LeeAnne Wang (Robert Kennedy lab).

Finally, nothing here would have been possible without my family, especially my mom Barbara, dad Vincent and brother Greg. My family instilled a love of science in me very early on and have encouraged me to keep pursuing academic studies, even in some uncertain times. Your endless support has allowed me to focus on grad school and achieve more than I thought was possible. Thank you.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

α-KG - α-ketoglutarate
6-dEB – 6-deoxyerythronolide B
A – adenylation
ACP – acyl carrier protein
ALAS – 5-aminolevulinic acid synthase
AONS – 8-amino-7-oxononanoate synthase
AOS - α-oxoamine synthase
AT – acyltransferase
ATP – adenosine triphosphate
BLAST – basic local alignment search tool
C – condensation domain
CoA – coenzyme A
CP – carrier protein
DH – dehydratase
ER – enoylreductase
FAD/FADH2 – flavin adenine dinucleotide/reduced flavin adenine dinucleotide
FAS – fatty acid synthase
GNAT – GCN5-related *N*-acetyltransferase
HPLC – high-performance liquid chromatography
hydride
ICP-MS – inductively coupled plasma mass spectrometry
KR – ketoreductase
KS – ketosynthase
LC – liquid chromatography
MS – mass spectrometry/mass spectrometer
MT or MTF – methyltransferase
NADP+/NADPH – nicotinamide adenine dinucleotide phosphate/nicotinamide adenine dinucleotide phosphate
Ni-NTA – nickel nitrilotriacetic acid
NRPS – nonribosomal peptide synthetase
Pant – pantetheine
PCP – peptidyl carrier protein
PE – partner enzyme
PKS – polyketide synthase
PLP – pyridoxal phosphate
Ppant – phosphopantetheine
PPTase – phosphopantetheinyltransferase
PSP – paralytic shellfish poisoning
SAH - *S*-adenosylhomocysteine

SAM – S-adenosylmethionine
SDM – site-directed mutagenesis
SNAC – *N*-acetylcysteamine
SPh – thiophenol
SPT – serine palmitoyltranferase
SSM – site-saturation mutagenesis
STX – saxitoxin
TE – thioesterase
TEII – type II thioesterase
UPLC – ultra performance liquid chromatography

Abstract

Natural products have driven advancements in many fields for over a century, from methods for structural elucidation to the development of new synthetic reactions to replicate their complex scaffolds and biological activities. Because enzymes perform challenging synthetic transformations with superior chemo-, site- and stereoselectivity, scientists have more recently been inspired to leverage Nature's biosynthetic machinery directly to make new compounds, establishing the emerging field of biocatalysis. This thesis describes studies with one broad class of proteins, carrier protein-dependent partner enzymes, to perform non-native reactions. Specifically, we developed biocatalytic platforms with two partner enzymes (PEs) SxtA AONS and AnaB to synthesize chiral amines chemoenzymatically.

PEs, found in the biosynthetic pathways of polyketides and nonribosomal peptides, catalyze a variety of reactions to assemble diverse and elaborate natural products. However, they are not often employed to mediate non-native reactions because of the high cost of correct substrate activation. By understanding how carrier proteins (CPs) can be used efficiently in tandem with PEs or how PEs may operate in the absence of CPs, PEs can be incorporated into the syntheses of valuable small molecules. Chapter 1 summarizes multiple approaches toward utilizing PEs for synthetic purposes and strategies for employing non-native substrates.

In Chapter 2, we characterized the native functions of the three PE domains within the polyketide-like synthase SxtA. We identified the correct starter unit and order of PE activity. Notably, the last domain, AONS transforms an amino acid into an α-amino ketone in a single step.

We envisioned using this enzyme as a general tool to derivatize amino acids with and without the CP.

We then optimized a CP-PE didomain platform with inexpensive acyl donors in Chapter 3 to perform the native reaction of SxtA AONS reaction scalably and to allow economical screening of the substrate scope. We observed ketone formation with multiple non-native thioesters and seven non-native amino acids but with very low overall conversion. We are improving the activity of SxtA AONS in a directed evolution campaign.

Chapter 4 describes our investigations in alpha-deuteration of amino acids, a CP-free reaction. Deuterated amino acids are valuable precursors toward labeled pharmaceutical agents but challenging to synthesize. SxtA AONS installs a deuterium atom on the α-carbon of select α-amino acids and all α-amino methyl esters assayed. Preparative-scale reactions allowed for stereoselective chemoenzymatic synthesis of an isotopically labeled analog of the drug safinamide.

Two SxtA domains, methyltransferase (MT) and AONS, were studied structurally and spectroscopically. In Chapter 5 we identify possible residues that lead to monomethylation in order to engineer additional dimethylation activity in SxtA MT. We also discuss UV-Vis studies with the AONS domain to understand the ketone-forming mechanism, elucidating the previously observed activity limitations and identifying possibilities for enzyme improvement.

Finally, Chapter 6 details our work with AnaB, a PE that operates on proline, the only proteinogenic substrate incompatible with SxtA AONS. AnaB natively oxidizes CP-bound proline to an iminium ion that may be stereoselectively functionalized with an exogenous nucleophile in the preparation of chiral cyclic amines. We sequentially built a four-step catalytic cycle around AnaB to use the CP efficiently.

The strategies presented in this thesis demonstrate the possibilities of leveraging CP-dependent PEs in chemoenzymatic synthesis. We anticipate that these approaches will also be applied to other PEs for the efficient, selective generation of chiral amines and other important synthetic building blocks.

Chapter 1: Introduction

With excerpts from "Carrier Protein-Dependent Biocatalysis". Chun, S. W.; Narayan, A. R. H. *Manuscript in preparation.*

Summary

With their potential to mediate reactions with nearly perfect chemo-, site- and stereoselectivity biocatalytic reactions and chemoenzymatic approaches have recently attracted significant attention and are now commonly integrated into industrial-scale processes. Carrier protein-dependent partner enzymes (PEs) from polyketide synthase and nonribosomal peptide synthetase pathways are one large class of biocatalysts that have untapped synthetic potential. To produce complex scaffolds, these enzymes function in coordinated assembly lines which could be exploited and tuned to afford custom molecules. However, PE reactivity in non-native reactions outside their biosynthetic contexts is not well-understood, and the requirement for the acyl substrates to be activated and covalently loaded onto small carrier proteins (CPs) is another expensive barrier to overcome. This Chapter reviews established methods to leverage PEs biocatalytically through both *in vivo* and *in vitro* strategies. Additionally, catalytic CP-PE platforms are proposed to enable more economical solution to the development of scalable methods using involving PEs.

1.1 Biocatalysis

Under the loosest definition of "biocatalysis", catalyzing chemical reactions with Nature's catalysts, humanity has been performing biocatalysis for millennia. In the past century, biocatalysis has evolved past simple fermentation of certain foods and alcoholic beverages to harnessing biocatalysts for chemical synthesis.[1,2] Early examples incorporated into industrial processes of commodity chemicals include chiral resolution of a racemic mixture by a lipase[3] and stereoselective synthesis of L-aspartic acid (**1.6**) from an achiral acid (Fig. 1.1A).[4,5] Biocatalysis has also found its way into the manufacturing of pharmaceutical products, such as the syntheses

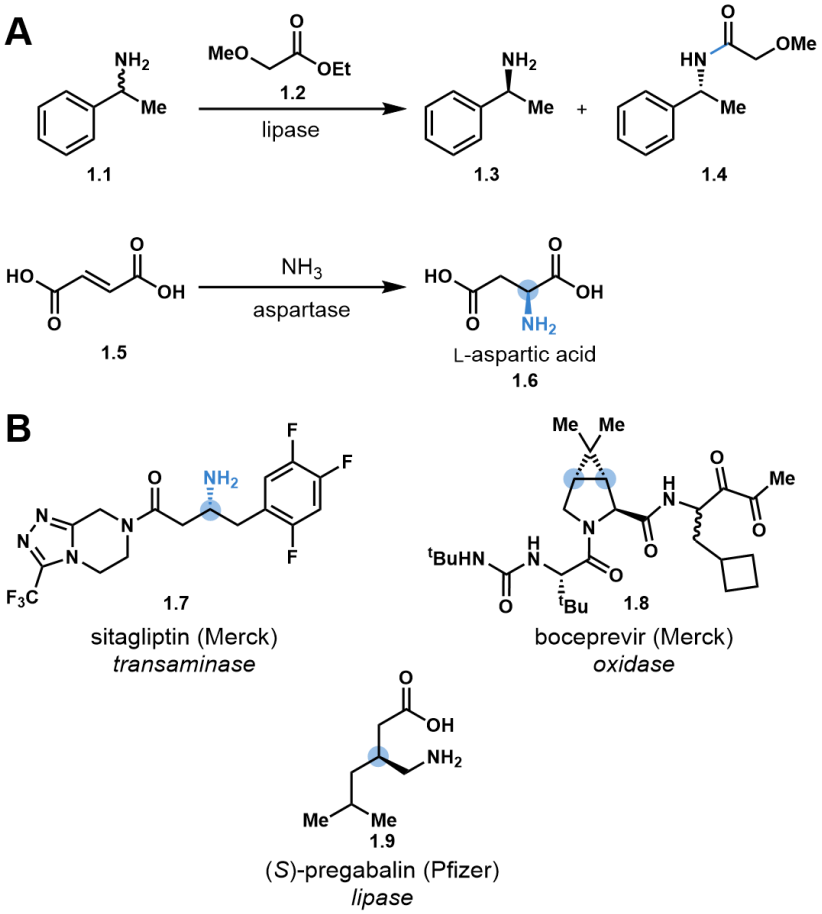


Figure 1.1. Commercially important biocatalytic reactions and products.
New bonds and stereocenters formed are highlighted in blue. (A) Selected examples of early biocatalytic transformations.[3,5] (B) Pharmaceutical products that have at least one enzymatic step in their industrial synthesis.[6–8]

of the compounds shown in Figure 1B.[6–8] Many of the first enzymes implemented into large-scale synthetic routes have focused on group interconversion reactions over more complex transformations like C–C bond formation.[2,9,10] Since then, enzymatic reactions have found increasing use based on several advantages over traditional chemical methods. Biocatalysts can mediate reactions with high chemo-, site- and stereoselectivity under mild conditions catalytically, resulting in relatively safe protocols that avoid the use of toxic and/or expensive reagents. One major challenge to wider incorporation of enzymes in chemical synthesis is in their substrate scopes.[1,11] Millions of years of selective pressure has provided many useful wild-type biocatalysts evolved to perform their native reactions and can therefore have limited activity on the non-native scaffolds required for a given synthetic route devised by a chemist.

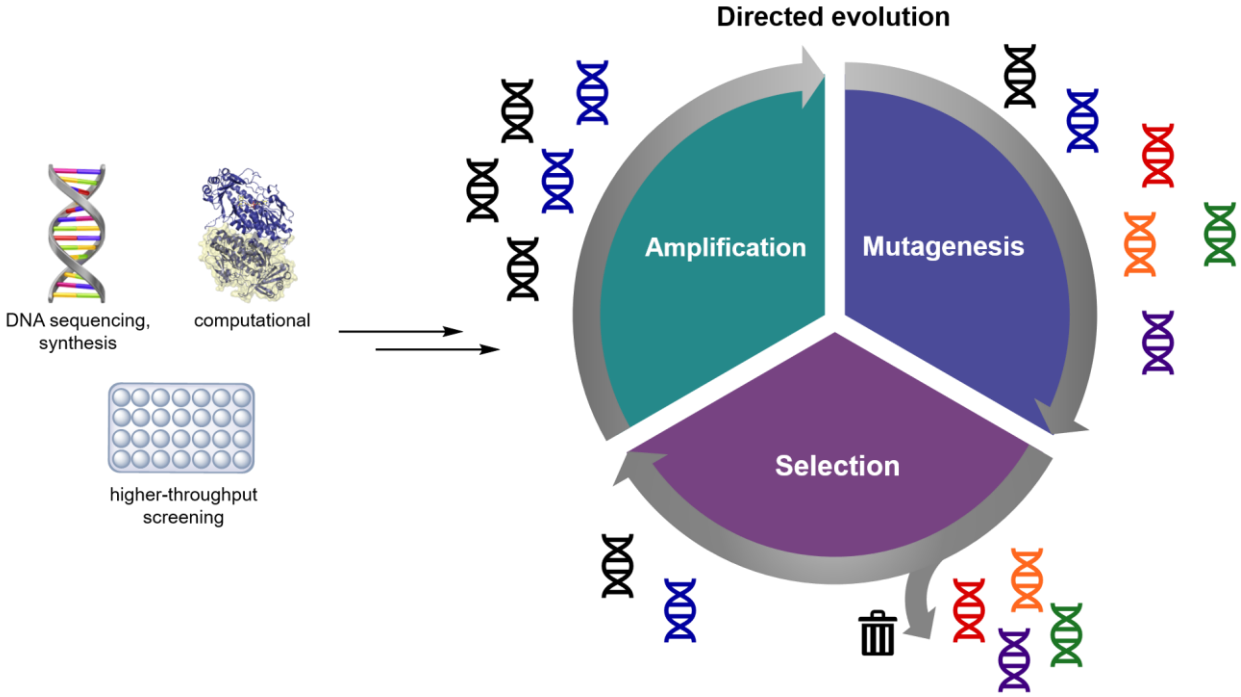


Figure 1.2. Advances in related fields enable improved directed evolution campaigns toward more active, robust and promiscuous biocatalysts.

Recent advances in several related fields have facilitated the implementation of biocatalysts through directed evolution for more sophisticated chemical transformations and synthetic routes

(Figure 1.2).[1] These developments include: (1) improved DNA sequencing and synthesis technologies for quickly providing more diverse parent sequences and faster mutagenesis;[12–15] (2) new higher-throughput assays to screen activity of protein variants;[16,17] (3) more computational power to model reaction pathways and predict useful mutations;[18–20] (4) rapid communication of results in online journals; [21,22] and (5) Prof. Frances Arnold winning the Nobel Prize in Chemistry in 2018 for directed evolution, which has spotlighted the field of biocatalysis.[1]

One broad group of enzymes that has experienced minimal use in preparative-scale reactions outside of *in vivo* processes is proteins from carrier protein-dependent pathways. Because of the relatively high cost of building blocks, large-scale carrier protein (CP)-dependent processes have largely been restricted to the fermentation of products such as 6-deoxyerythronolide B[23] and lovastatin (see section 1.2.1).[24]

CP-dependent pathways include fatty acid synthases (FASs), types I and II polyketide synthases (PKSs), and nonribosomal peptide synthetases (NRPSs) that biosynthesize valuable natural products such as antibiotics erythromycin, vancomycin or the siderophore enterobactin (Figure 1.3A).[25,26] Endogenously, the partner enzymes (PEs) in these pathways operate on substrates that are covalently bound to carrier proteins (CPs) and not free molecules in solution (Figure 1.3B). Remarkably, CPs across all types of CP-dependent pathways share a conserved three-helix structure despite low protein sequence identities across this class of proteins. CPs also possess a conserved serine residue, to which a 4'-phosphopantetheine (Ppant) cofactor is post-translationally attached to activate the CP from the *apo* to *holo* form (**1.16** to **1.17**, Figure 1.3C). From here, substrates covalently tethered through a thioester linkage to the thiol of the Ppant arm, and then delivered to PE active sites (Figure 1.3D).[27,28]

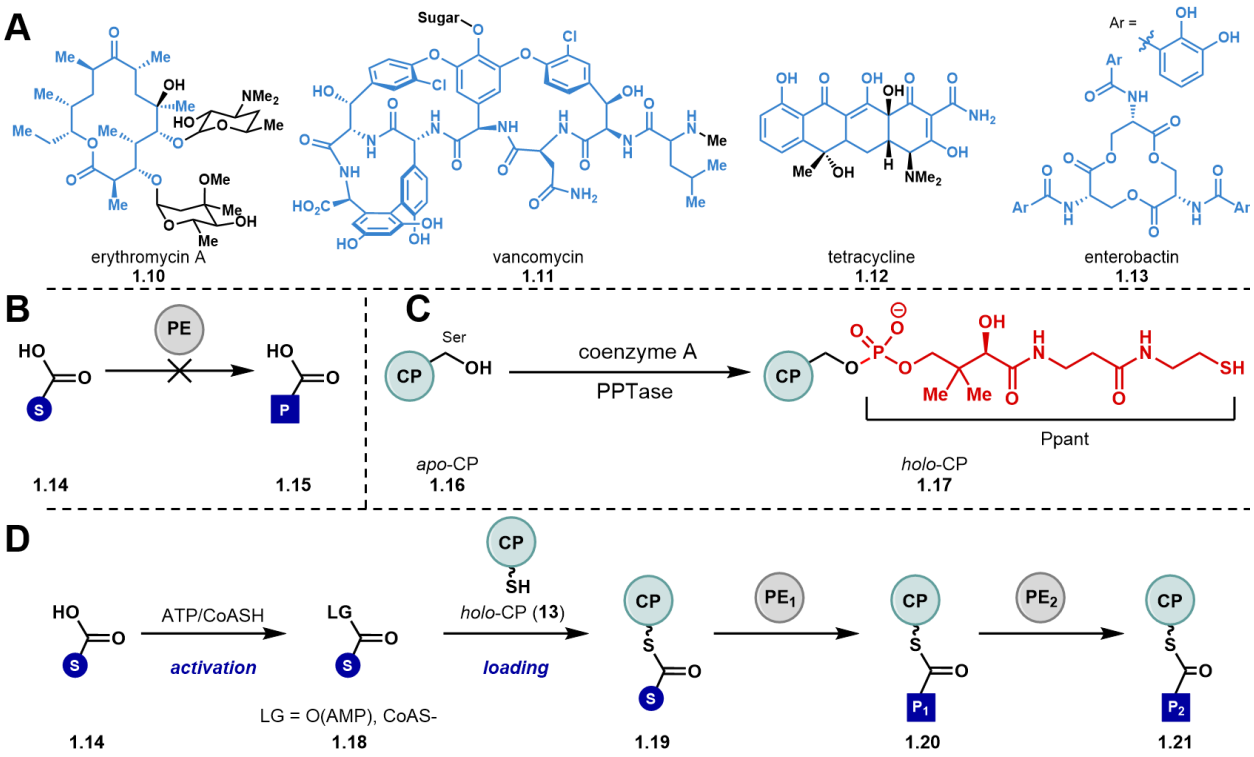


Figure 1.3. CP-dependent PEs assemble complex natural products via intermediates covalently bound to carrier proteins.
(A) Some natural products resulting from CP-dependent pathways. Portions assembled by PEs are highlighted in blue, and the black moieties are modified in post-CP segments. (B) PEs are inactive on free acids. (C) Post-translational modification of inactive apo-CPs by phosphopantetheinyl transferases (PPTases) to form the active *holo*-protein. (D) General reaction cycle of substrate activation into esters or thioesters, loading onto *holo*-CPs and modification by CPs.

CP-dependent pathways also are found in interesting architectural arrangements. In type I FAS, PKS and NRPS pathways, multiple CPs and PEs are expressed as domains in one polypeptide chain to form megasynthases, while these elements are expressed as discrete proteins in type II systems.[27–29] Type III PKSs also exist but are CP-independent pathways that use solely coenzyme A (CoA) thioesters, and thus will not be discussed here.[30] Types I and II CP-dependent pathways are further organized into modules, each of which is responsible for one cycle of chain initiation or elongation (Figure 1.4). Following the selection of the starter unit by the loading module, additional modules elongate the acyl or peptidyl chain by incorporating extender units and

optionally elaborate on that unit with tailoring PEs. Chain extension can continue in multiple cycles until the chain is terminated by one of many offloading mechanisms such as a thioesterase (see Du and Lou[31] for a comprehensive overview). Further modifications by post-PKS or -NRPS enzymes are also possible (black moieties in Figure 1.3A).

Because most type I pathways react in a predictable assembly line-fashion from the N- to the C-termini of each multi-domain protein, scientists have long been interested in designing pathways to synthesize custom natural product derivatives.[32–34] The simplest approach of splicing together the domains of interest has been met with very limited success, as it has become clear that the precise timing of PE reactivity is tightly controlled by linkers and complex protein-protein interactions between domains and modules.[35–40] The interactions themselves are difficult to characterize structurally for multiple reasons (limited protein stability, CPs are small and flexible, megasynthases are typically too large for x-ray crystallography), so we and many other groups have focused on studying individual or a small subset of domains rather than full modules. We anticipate that our studies will contribute to the body of literature on PEs operating on non-native substrates in order to develop syntheses of novel small molecules by PEs.

The remainder of this Chapter will discuss previously published examples of utilizing PEs in biocatalytic reactions, and the basis of our proposed CP-PE platform. I will then briefly summarize the contents of research Chapters 2-6.

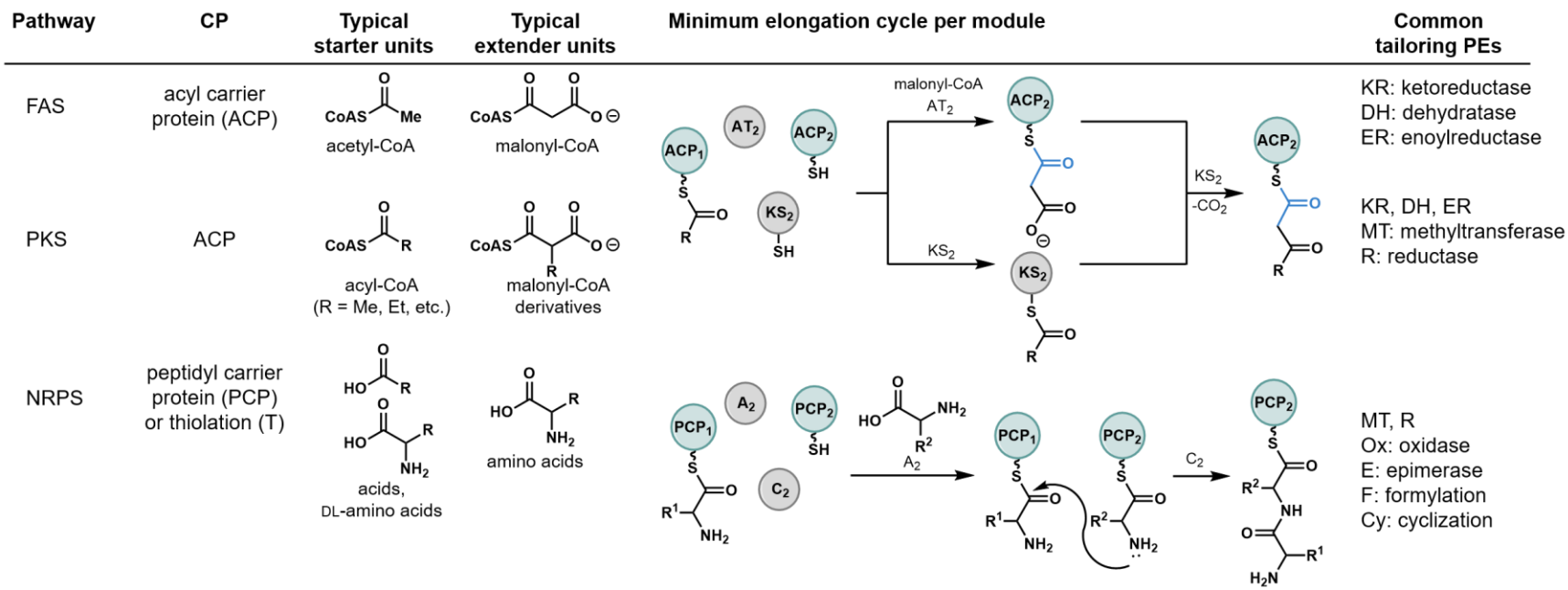


Figure 1.4. Common features of fatty acid synthase (FAS), polyketide synthase (PKS) and nonribosomal peptide synthetase (NRPS) pathways.

Minimum elongating components are acyltransferase (AT), ketosynthase (KS), ACP in PKS and adenylation (A), condensation (C), and PCP in NRPS.

1.2 Summary of Established CP-PE Biocatalytic Approaches

CP-dependent pathways are relatively expensive in their native contexts, requiring ATP and/or CoASH for activation of starter and extender units (Figure 1.5A), and the synthesis of sometimes very large proteins.[28] Even so, synthetic biologists, biochemists and chemists remain fascinated by these pathways because they are responsible for the biosyntheses of complex commercially valuable scaffolds,[41] because CP-dependent PEs can catalyze chemically challenging transformations, and because of the future possibility of merging reactions into cascades.[34]

Some examples of PE-catalyzed transformations are shown in Figures 1.5B-D. After the ketosynthase-catalyzed condensation in PKS proteins (Figure 1.4), PKS ketoreductases (KR) reduce the β-keto groups to β-hydroxyl groups stereospecifically with NADPH (nicotinamide adenine dinucleotide phosphate hydride), a reducing cofactor. Substituted carbons alpha to the thioester are not configurationally stable due to the low pK_a; thus, the reduction of β-carbonyls sets both α- and β-stereocenters simultaneously.[42] KR subtypes that catalyze formation of all four combinations of stereocenters have been discovered (Figure 1.5B).[43] Optional loss of water by dehydratases (DH) may follow the KR-catalyzed reaction, forming α,β or β,γ unsaturated bonds. GphF DH1, the first DH in the gephyronic acid biosynthetic pathway, initially forms an α,β-unsaturated intermediate **1.30** from **1.29** (Figure 1.5B), but then isomerizes **1.30** so that the final product **1.31** contains a β,γ double bond and the α-stereocenter is epimerized from its original configuration.[44] A third PE-catalyzed reaction, that is challenging to accomplish with traditional chemical methods is the cyclopropane ring formation in kutzneride biosynthesis.[45] Chemical cyclopropanation reactions often employ extremely reactive intermediates such as carbenes or carbenoids,[46] whereas the ring formation is accomplished biosynthetically by a two-enzyme

cascade under much milder conditions. An isoleucyl substrate bound to the PCP KtzC **1.32** is first chlorinated by the halogenase KtzD and then cyclized by KtzA, eliminating HCl (Figure 1.5D).[45,47]

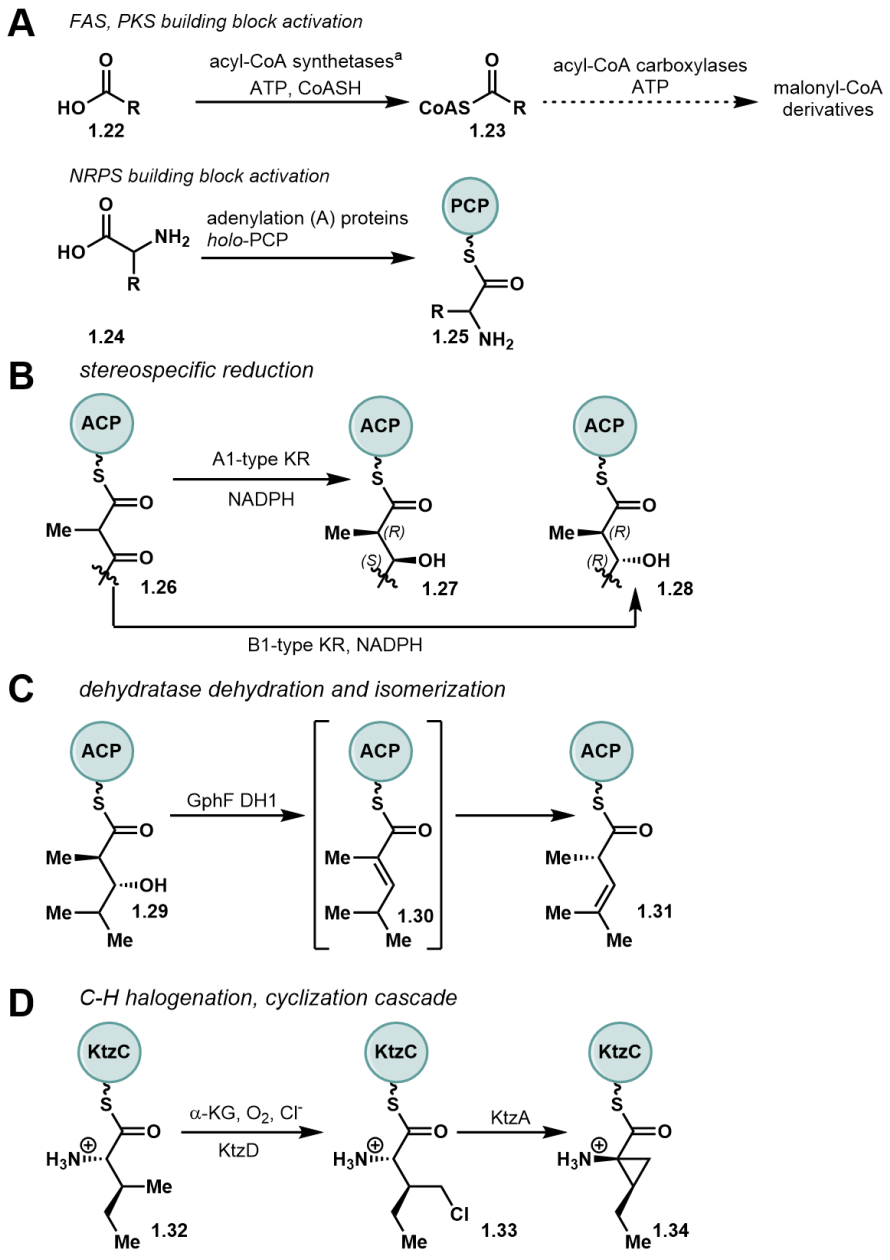


Figure 1.5. Selected transformations performed by CP-dependent enzymes.
(A) Preparation of starter and extender units. (B) Stereospecific β-keto reduction by ketoreductases (KRs). (C) Dehydratase (DH) elimination of β-hydroxy groups, with optional double bond isomeration. (D) Chlorination of an unactivated C-H bond by the α-ketoglutarate (α-KG) non-heme iron halogenase KtzD, followed by a KtzA-catalyzed cyclization.[43,44,47]

Using natural cascades such as KtzACD or a designed non-native sequence would be ideal in synthetic route development; however, many challenges remain to render CP-dependent PEs amenable to preparative-scale synthesis, including the aforementioned protein instability, required CP post-translational modification and activation cost to form the required CP-bound intermediates. Chemists and biochemists have developed multiple methods to mitigate the last issue that will be covered here. This Chapter is organized according to strategies using natural producer strains *in vivo*, using native loading mechanisms *in vitro*, bypassing CP-dependent gatekeepers nonenzymatically, and finally engineering a PE away from CP dependence.

1.2.1. CP-dependent pathways *in vivo*

Prior to recent advancements in genome sequencing, natural products were discovered and elucidated by isolation from bacterial and fungal extracts of field samples or laboratory-grown cultures.[48] While *in silico* genomic mining is preferred today for polyketide and nonribosomal peptide discovery,[49] *in vivo* methods involving the culturing of natural producers are also used for industrial-scale polyketide production and analytical feeding studies to identify a product's fundamental building blocks.[50]

These methods are particularly useful when the biosynthetic gene cluster encoding a target pathway has not yet been located.[50] Even if a gene cluster and its associated natural product(s) are known, culturing the producing organisms avoids the challenges associated with cloning and recombinant expression of megasynthases for *in vitro* experiments.[51] However, employing *in vivo* approaches requires producing strains to be culturable and the gene cluster of interest to be expressed under laboratory conditions. This section will discuss examples of culturable strains for

commercial fermentation to manufacture polyketide and nonribosomal peptide-based products, and smaller-scale instances of native product analogs generated through feeding experiments.

1.2.1A. Commercial fermentation

In order to supply the billions of dollars' worth of commercial polyketide and nonribosomal peptide products, industrial processes obtain valuable compounds such as erythromycin A (**1.10**, see Figure 1.3A) by large-scale fermentation of optimized producing organisms followed by isolation of the desired molecules.[51] The compounds may be the final product of a full PKS or NRPS pathway,[51] or for semisynthetic routes, an advanced intermediate that can be further tailored chemically.[52] Maintaining a constant supply of the required activated starter and extender units and reducing equivalents for a CP-dependent pathway is expensive, motivating industrial-scale processes to rely on the optimized host strains to produce all required building blocks and perform optional difficult post-CP tailoring.[53,54] To the best of our knowledge, large-scale fermentation is the most cost-effective method to obtain large amounts of compounds from primarily CP-dependent pathways but has only been applied to use the native reactions of PEs. A selection of well-known industrially produced polyketide and nonribosomal peptide products obtained from natural producers is listed in Table 1.1, with product titers ranging from milligram to multigram quantities per liter of culture. Many of these entries include the post-PKS/NRPS modifications as information on the yields of only the CP-dependent portions of pathways were not available.

Although impressive yields have been reported for products such as oxytetracycline and vancomycin (entries 10 and 12),[55,56] work to improve industrial strains and product yields are ongoing.[57] Typically, industrial strains have been and continue to be optimized over decades by random mutagenesis;[58] therefore, many of these titer values are likely outdated. Recently, efforts

to further improve yields by metabolic engineering to increase protein expression and availability of building blocks while decreasing competition for precursors, shunt product formation or product degradation have been reported, with promising results.[59] Some fungal hosts are challenging to genetically manipulate, and there is interest in transferring the pathways to *E. coli* that are simpler to culture and manipulate. However, these *E. coli* strains (entries 1 and 2) are not yet as productive as the native hosts.[60,61] One notable example is the *E. coli* BAP1 strain originally developed by Pfeifer and Khosla in 2001 for the biosynthesis of 6-deoxyerythronolide B (6-dEB).[62] Due to the promiscuous phosphopantetheinyl transferase (PPTase) gene *sfp* inserted into its genome, the BAP1 strain has found widespread use as a host to prepare recombinant *holo*-carrier proteins today.

Table 1.1. Commercial products from CP-dependent pathways obtained by industrial-scale fermentation

Entry	Isolated product	Pathway	Strain	Product titer (per L of culture)
1	6-dEB (PKS product)	Erythromycin A	*E. coli* BAP1	100-1100 mg[23,62,63]
2	Erythromycin A (final pathway product)		*E. coli* LF01	10 mg[54,64]
3	Erythromycin A		*Saccharopolyspora erythraea*	8-10 g[51]
4	Tylosin	Tylosin	*Streptomyces fradiae* TM-224	7.2 g[65]
5	Lovastatin	Lovastatin	*Aspergillus terreus* LA414	660 mg[24]
6	Penicillins G and V	Penicillins	*Penicillium chrysogenum* and *P. rubens*	50 g[66]
7	Cephalosporin C	Cephalosporin	*Acremonium chrysogenum*	3.1 g[67,68]
8	Daunorubicin	Doxorubicin	*Streptomyces peucetius*	60 mg[69]
9	Tetracycline	Tetracyclines	*Streptomyces lusitanus*	11.1 g[70]
10	Oxytetracycline		*S. rimosus* 23833	55 g[55]
11	Chlortetracycline		*S. aureofaciens* F3	14.8 g[71]
12	Vancomycin	Vancomycin	*Amycolaptosis orientalis* KCCM-10836P	11.5 g[56]

Currently, fermentation is the only scalable method than generates orders of magnitude more PKS/NRPS natural products than industrial fermentation. Until significant advancements in engineering how PEs utilize substrate surrogates are discovered, fermentation will likely be the only economical strategy for lengthy polyketide and nonribosomal peptide production.

1.2.1B. Feeding studies and precursor-directed biosynthesis

On laboratory-scale, polyketide and nonribosomal peptides are routinely isolated from much smaller scale cultures. In feeding experiments, the culturable host's biosynthetic pathways still generate and activate most of the building blocks, but supplementing culture media with isotopically labeled substrates can lead to incorporation into natural products. This has been used in all types of biosynthetic pathways for decades to track metabolic processes, including classic feeding studies in the 1950s and 1960s to provide clues on the fundamentals of CP-dependent biosynthesis.[72,73] Isotopic feeding experiments are still employed to identify PKS and NRPS extender units, and elucidate structures of the small molecule products.[74,75]

Once the precursors are known, their analogs may also be added to cultures of producing strains in a process known as precursor-directed biosynthesis (PDB). Provided the biosynthetic enzymes are sufficiently promiscuous, the compounds isolated are new product analogs.

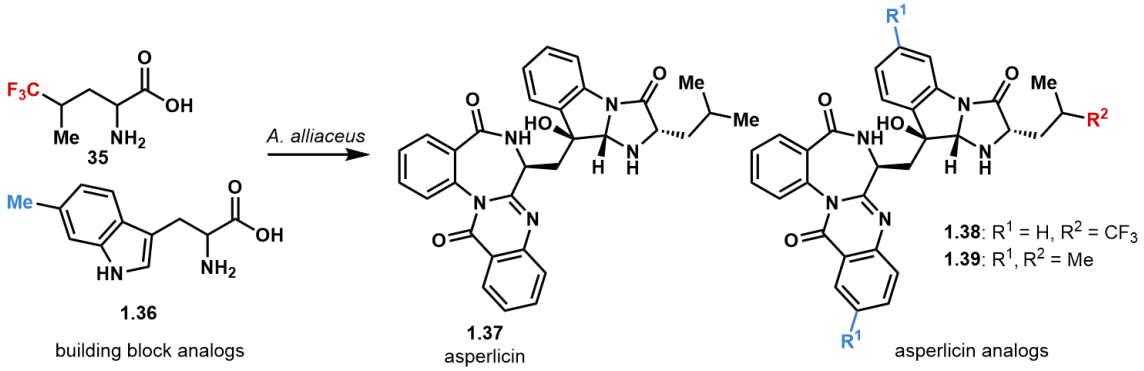


Figure 1.6. Precursor-directed biosynthesis of asperlicin analogs from leucine and tryptophan derivatives by the asperlicin producer *Aspergillus alliaceus*.
See Houck et al.[76.]

For example, fluorinated leucine (**1.35**) or methylated tryptophan (**1.36**) added to the media of asperlicin producer *Aspergillus alliaceus* gave novel asperlicin analogs such as **1.38** and **1.39** (Figure 1.6).[76] PDB has been used to produce new analogs of cyclosporins,[77] among others.

Analogs of the balhimycin[78], salinosporamide,[79] and many other natural products have been synthesized by mutasynthesis,[38] a variation of PDB using mutant or auxotrophic host strains.

1.2.2. *In vitro* strategies with native loading methods

PDB and mutasynthesis require potential building blocks to be stable and taken into host cells, properties that all desired precursors may not possess. Additionally, recently published genome sequences of potential natural product producers have revealed the presence of many silent biosynthetic gene clusters that are not expressed under normal laboratory conditions.[48] Fortunately, advances in cloning and heterologous protein expression now permit more direct probing of CP-dependent pathways.[80] For *in vitro* reactions with purified enzymes, studying PEs in a simplified system without competing host reactions can provide advantages, but the gene cluster must be known before this can be used as a viable approach.[81] However, unlike the *in vivo* processes described in Section 1.2.1 where the host organism pathways activate building block precursors, employing native loading machinery *in vitro* also requires the addition of activating cofactors. The most common loading PEs, NRPS adenylation proteins (A) and PKS acyltransferases (AT), require ATP and CoASH, respectively, which may be prohibitively expensive at larger scales.[82]

1.2.2A. Partial pathway reconstruction with NRPS adenylation domains

Whereas the *in vivo* strategies discussed previously (Section 1.2.1) primarily use complete pathways, smaller subsets of PEs have been probed through *in vitro* experiments. If an interesting PE is found within the first module of a CP-dependent pathway, one direct method of generating loaded CP substrates *in situ* is through the native loading domains.[83] For example, to probe the halogenation activity of PltA, Dorrestein and coworkers used a three-enzyme cascade to first

adenylate and load the starter unit L-proline (**1.40**) onto the *holo*-PCP PltL, which is subsequently oxidized from pyrrolidine **1.41** to pyrrole (**1.42**) by PltE (Figure 1.7A).[84] PltA was then demonstrated to incorporate both bromide ions in addition to its native chlorination activity.[85] In another report, Mann and coworkers examined the substrate scope of a PltE-related oxidase, AnaB, by adenylating and loading the native L-Pro substrate as well as substituted derivatives onto the cognate PCP.[86,87] Stachelhaus and Walsh studied a longer sequence involving the putative

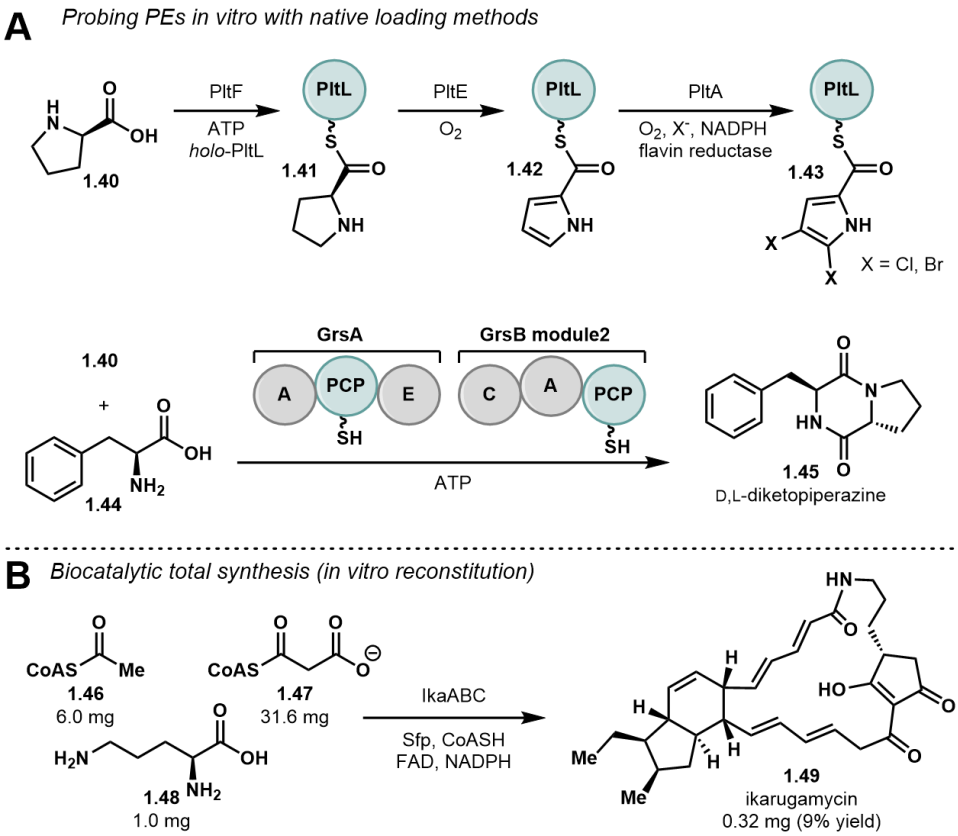


Figure 1.7. *In vitro* PE studies with native loading methods.
(A) Probing the initial steps of pyoluteorin biosynthesis – loading of the L-Pro (1.40) starter unit onto PltL, followed by a four-electron oxidation and chlorination.[84,85] (B) The epimerase domain of the gramicidin S initiation module GrsA epimerizes L-Phe (1.44) to the D-enantiomer.[88] (C) One-pot biocatalytic total synthesis of ikarugamycin acyl-CoA and L-ornithine building blocks.[89] A: adenylation; E: epimerase; FAD: flavin adenine dinucleotide.

epimerase domain of GrsA (Figure 1.7A) and condensation by the next module by incubating GrsA, GrsB module 2, ATP, L-Phe (**1.44**) and L-Pro (**1.40**) together.[88] They found that the

dipeptide was released spontaneously, forming an L-Phe- D-Pro diketopiperazine (**1.45**) that was verified by comparison to enantiopure authentic standards.

1.2.2B. Biocatalytic total synthesis

Biocatalytic total synthesis (sometimes known as *in vitro* reconstitution or total biosynthesis) with the goal of isolating and characterizing the fully assembled product, is a greater challenge than the partial pathway studies described in the previous section. Analogous to chemical total synthesis, this method aims to completely build a target molecule from fundamental building blocks but is possible to perform in one pot. Among the earliest biocatalytic total syntheses for PKS were tetracenomycin[90] and actinorhodin,[91,92] and more recent reports have been published on the total biocatalytic synthesis of RK-682,[93] cladosporin[94] and one example of 6-deoxyerythronolide (6-dEB) where methylmalonyl-CoA building blocks could be substituted by ethylmalonyl-CoA to biosynthesize new 6-dEB analogs.[95] Reconstitutions have been reported for NRPS pathways of malbrancheamide,[96] asperlicins C and D[97] and vibriobactin,[98] among others. The ikarugamycin gene cluster contains IkaA, a hybrid PKS-NRPS protein, and IkaBC, two-post PKS tailoring enzymes.[89] Incubation of the three proteins with acyl-CoAs and L-ornithine (**1.46**-**1.48**, Figure 1.7B), the PPTase Sfp and required cofactors in one pot afforded the final product ikarugamycin in a 9% isolated yield. Although not quantitative, it is a significant improvement over the most efficient total synthesis reported to date requiring a linear sequence of 27 steps to afford an overall yield of 0.11%.[99,100] The cost of the biocatalytic reaction was further reduced by adding the CoA ligase MatB to biosynthesize the acyl-CoA units *in situ* from free acids, ATP and CoASH, which also decreased the CoASH to catalytic amounts.

In a more dramatic example, Moore and coworkers biosynthesized the polyketides enterocin and wailupemycin *in vitro* from a mixture of twelve proteins, malonyl-CoA, benzoic acid, ATP, *S*-adenosylmethionine (SAM) and NADPH for a final yield of 25%.[101] The chemical total synthesis of enterocin has not been accomplished yet.[102] For more exhaustive lists of biocatalytic total syntheses in CP-dependent systems, detailed reviews have been published.[103,104]

Despite the advantages of these biocatalytic methods with native loading enzymes over traditional synthesis, the activation cost by ATP and CoASH may be prohibitive, particularly when utilizing less efficient non-native substrates. Other cost-effective chemoenzymatic and nonenzymatic methods have also been developed in parallel to these strategies to study CP-dependent pathways (see Section 1.2.3).

1.2.3. Chemoenzymatic methods

1.2.3A. Substrate surrogates

Exploiting traditional chemical methods to study PEs can also be an effective strategy, providing opportunities including: (1) lower costs for preparation of substrate surrogates, (2) investigation of PEs on advanced intermediates, and (3) utilization of non-native substrates.

The CP attachment is not absolutely required for all PEs, as some enzymes catalyze reactions on free substrates activated by a selection of surrogates that are far less complex than a CP (Figure 1.8A). Generally, these mimic the Ppant arm of *holo*-CPs (**1.13**) or use CoASH-derived thioesters.[105] PEs such as *E. coli* BioF[106] have been reported to operate on CoA thioesters as well, whereas the squalestatin tetraketide synthase (SQTKS) enoylreductase (ER) reduces a shorter panthetheine thioester stereoselectively (**1.57**, Figure 1.8B).[107] A simplified thiol, *N*-acetylcysteamine (SNAC), was originally used to study FAS systems and is now the most common

surrogate.[105,108] Many PEs mediate reactions directly on SNAC thioesters, as Keatinge-Clay reported for six related excised *C*-methyltransferase (*C*-MT) domains that methylated 1,3-dicarbonyl substrates at the alpha position (**1.59** to **1.60**).[109] Typically, NRPS condensation (C) domains catalyze peptide bond formation of a growing nonribosomal peptide from two aminoacyl-PCP substrates, but Ding and coworkers demonstrated that one of the two CP-bound substrates could be replaced by the corresponding SNAC thioester.[110] The C domain of cryptophycin biosynthesis module 6 (CrpD) formed an unusual ester bond between an PCP-bound α-hydroxy extender unit and an SNAC analog of the module 5 product. Other examples of free thioester substrates have been published for CP-dependent ketoreductases (KR),[43] dehydratases (DH)[111], oxidases[112] and additional PEs.[113,114]

Substrate surrogates have a second mode of action: ketosynthases (KS) and thioesterases (TE) may mediate their respective reactions after surrogate molecules have transacylated PE catalytic residues (Cys or Ser, Figure 1.8C). In many reports, surrogate molecules acylate a KS active site Cys residue while its cognate CP is enzymatically loaded with an extender unit. The KS can then catalyze the expected chain extension.[115–117] In one example, the Khosla and Cane groups acylated the sixth and final KS of 6-dEB synthase 3 (DEBS 3) with non-native diketide SNAC **1.61**, while the associated AT domain loaded ACP_6 from a methylmalonyl-SNAC (**1.62**) extender unit source.[118] After KS-mediated extension and further reduction by KR_6, the DEBS TE domain cyclized the triketide intermediate **1.64** to lactone **1.65**. In addition to their inexpensive preparation relative to the corresponding CoA thioesters (methylmalonyl-CoA was 4000 times more expensive per mole than methylmalonyl-SNAC **1.62** in 1998),[118] these small molecules are employed in whole-cell applications as they are frequently cell-permeable, whereas CoA compounds and CPs are not.[119]

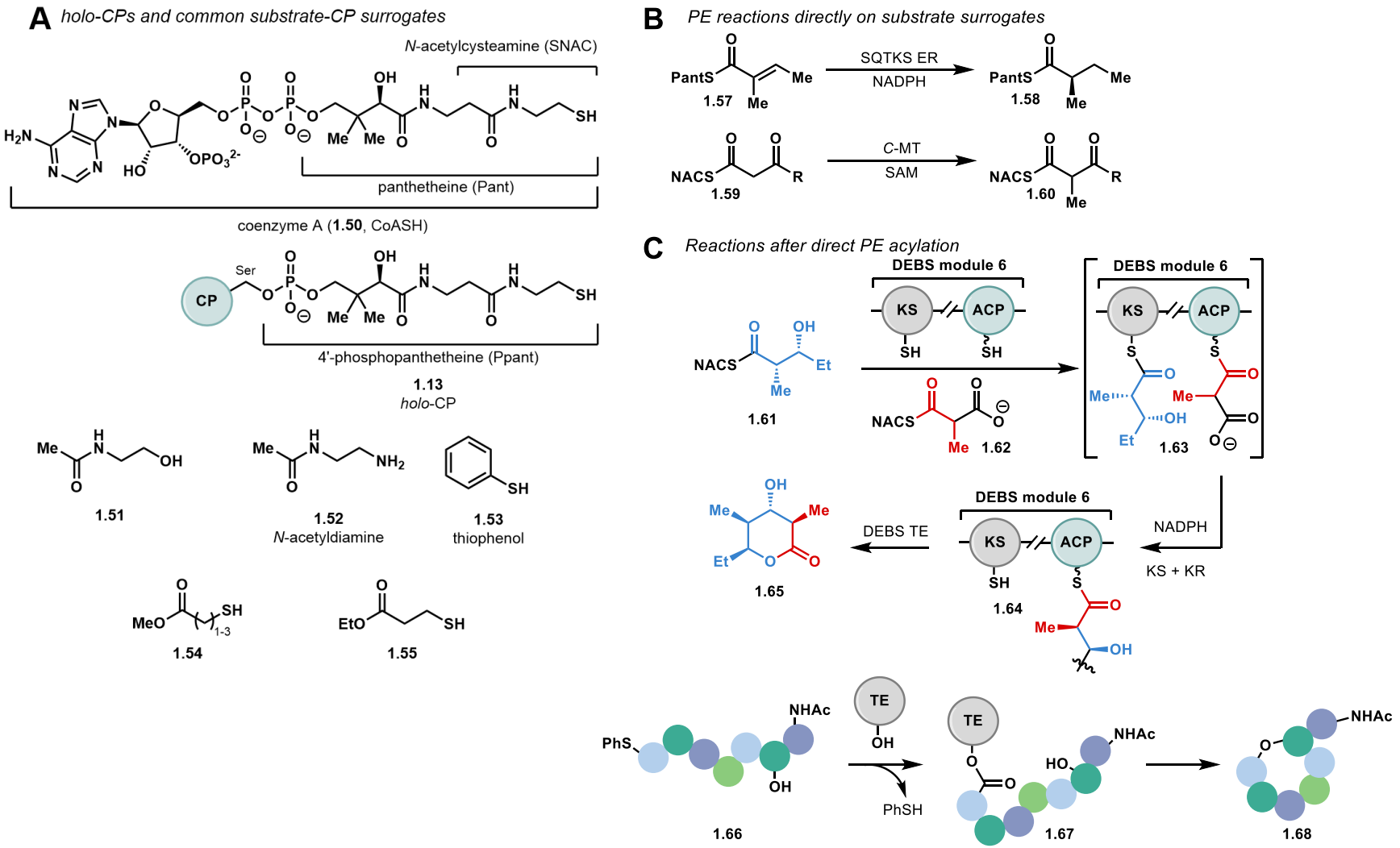


Figure 1.8. Chemoenzymatic PE methods with substrate surrogates and acyl donors.
(A) Comparison of *holo*-CP and common CP surrogate structures.[105] (B) Selected examples of PE activity on free surrogates.[107,109] (C) Selected examples of PE activity after transacylation by substrate surrogates.[118,120] SQTKS ER: squalestatin tetraketide synthase enoylreductase domain; SAM: *S*-adenosylmethionine; DEBS: 6-deoxyerythronolide B synthase; KS: ketosynthase; TE: thioesterase.

Although substrate surrogates are a widely used tool to probe late-stage PEs, it is unfortunately not yet possible to predict whether a PE is compatible with surrogates or predict the PEs for which the protein-protein interactions with their cognate CP are critical. A number of reports disclose PE and free substrate incompatibility,[121–123] and it is likely that many other similar negative results have gone unpublished. Additionally, PE activity may be dependent on the identity of surrogate activating group. Sieber and coworkers observed that fengycin TE was not active with an SNAC thioester source, but the active site Ser residue could be acylated by the more reactive thiophenol esters (**1.66**, see Figure 1.8C), leading to the desired macrocylized products.[120] When incubated with the KS_6-AT_6-ACP_6 tridomain in PikAIV and TE PikAV, Hansen and coworkers reported that the distribution of hexaketide substrates directly cyclized by the TE or that acylated the KS for elongation to a heptaketide and later cyclization shifted dramatically among the five activating groups investigated.[124]

1.2.3B. Generation of acyl-CP substrates by non-native enzymatic methods

For completely-CP dependent systems, there are both enzymatic and nonenzymatic methods reported to prepare acyl-CP substrates. Enzymatic methods utilize the post-translational modification necessary for native CP activity: promiscuous PPTases and CoA coupled to a desired substrate can be incubated to prime and load a CP in a single step (Figure 1.9A).[125] The most commonly used PPTases are Svp from *Streptomyces vertilicus* and Sfp *Bacillus subtilis*, which are compatible with many ACP, PCP and CoA derivative combinations.[126–128]

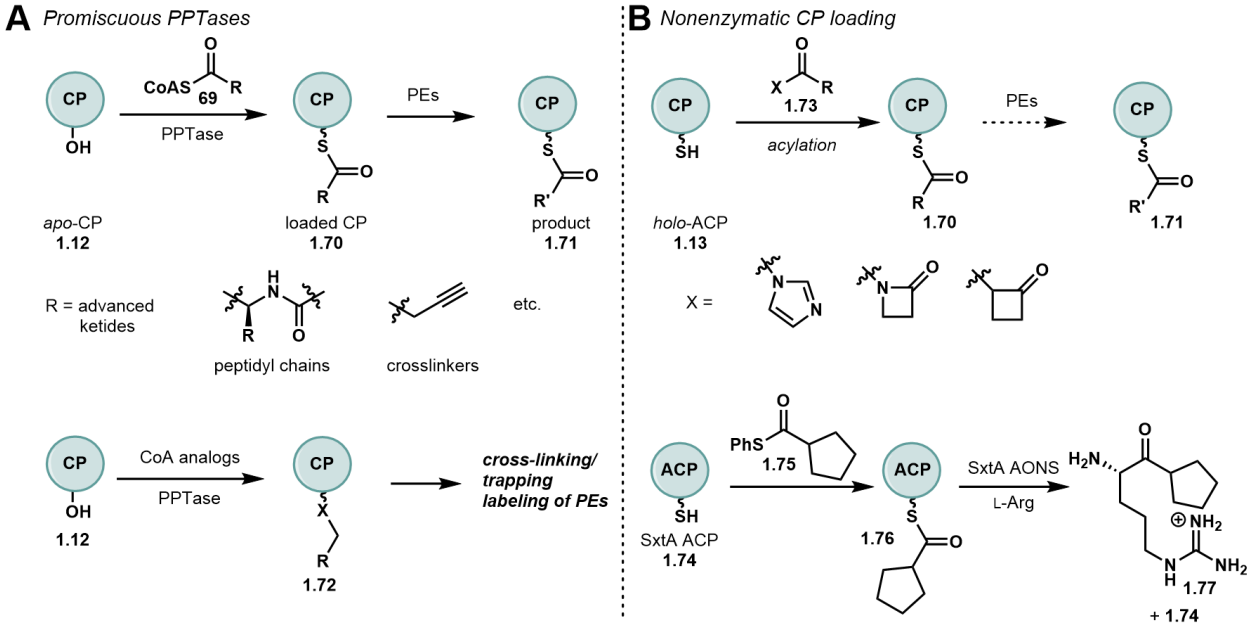


Figure 1.9. Chemoenzymatic methods to generate loaded CPs.
(A) Enzymatic loading via promiscuous PPTases from *apo*-CP.[125] (B) Nonenzymatic loading by small molecules.[129–131]

The acyl groups may be putative starter units,[132] non-native substrates,[133] advanced ketides[134,135] or oligopeptides,[136] among others.[125] These substrates are used to bypass the native loading mechanisms that limit the incorporation non-native building blocks and to examine PE substrate scopes, especially in later modules. The Burkart group has done extensive research with highly modified CoA derivatives, employing fluorescent, non-thioester and non-hydrolyzable

CoA-based compounds to identify CP active sites or capture interactions between a PE and a CP loaded with a cross-linking substrate (Figure 1.9A).[137] The latter has been utilized to elucidate PE mechanisms and helped to solve new structures of CP-dependent enzymes.[125,138]

CoA derivatives are typically prepared by traditional organic synthetic strategies or oligopeptides by solid-phase peptide synthesis, followed by coupling to CoASH or an analog.[139] The cost can be reduced further by coupling to a pantetheinyl derivative (Figure 1.9A) and generating the CoA derivative *in situ* through the CoA biosynthetic enzymes CoaADE.[140] Any of the above uses of acyl-CPs to study PE activity without a method to regenerate *apo*-CP are limited to one CP turnover.

1.2.3C. Generation of acyl-CP intermediates by nonenzymatic methods

As a cheaper alternative to complex CoA derivatives, acyl-CPs may be generated nonenzymatically through thioesterification and transthioesterification reactions between the reactive thiol of a Ppant arm and activated acyl substrates. These methods only react with *holo*-CP, requiring recombinant *apo*-CPs to be primed *in vitro* if the *in vivo* post-translational modification is incomplete.[128]

First reported in 1981 for loading FAS ACPs, *N*-acylimidazolides (**1.73**, Figure 1.9B) have also been used to nonenzymatically acylate oxytetracycline and griseusin PKS ACPs.[141,142] Crosby and coworkers found that the imidazolides also reacted with exposed cysteine residues, necessitating a Cys to Ser mutation in ACP substrates before Ppant-specific labeling occurred. The Schnarr group has explored β-lactams and -lactones as activating groups and similarly observed competition between the Ppant thiol and other reactive nucleophilic residues, notably KS catalytic cysteines.[129,130] As described in section 1.2.3A, activated small molecule surrogates react

preferentially with KS domains over ACPs in PKS systems.[143] Likewise, when Sieber attempted to load PCP-TE didomains from aminoacyl-thiophenol esters, only acylation of the TE domain was observed (see Figure 1.9B).[120] Thus, until extremely selective activating groups are discovered, diffusive loading is limited to systems without competition from other exposed residues that can be nonspecifically labeled. Our group reported one example recently for the non-KS-containing polyketide-like synthase SxtA, where the excised ACP was loaded by a thiophenol-activated substrate. The resulting acyl-ACP could then react with a condensing PE to release a small molecule product (**1.77**).[129,130131]

1.2.4. Directed evolution to reduce CP-dependence

For PEs that already accept small molecule substrate surrogates, perhaps the most economical strategy for their use may be an engineered variant possessing enhanced activity in the absence of CPs. Tang and coworkers have demonstrated this once in LovD. The wild-type enzyme is an acyltransferase that natively transfers a LovF ACP-activated acyl group (**1.78**) monacolin J acid (**1.80**) to produce lovastatin precursor **1.81** (Figure 1.10).[144] LovD also accepts an α-dimethylbutyryl substrate surrogate (**1.82**) as well that leads to a precursor (**1.84**) of simvastatin,[145] which is more active than statin **1.81**.

The native LovF-based reaction is reported to be 1300x more active than the surrogate reaction but through directed evolution, LovD was evolved to increase protein expression levels and stability. the wild-type LovD-LovF transformation remains superior, the final seventh generation LovD Each successive generation was more active toward **1.82** and less active with acyl-LovF **1.78**. The final variant showed 11-fold whole-cell activity improvement and converted >30 g/L of monacolin J acid **1.84** to simvastatin acid per day.[146]

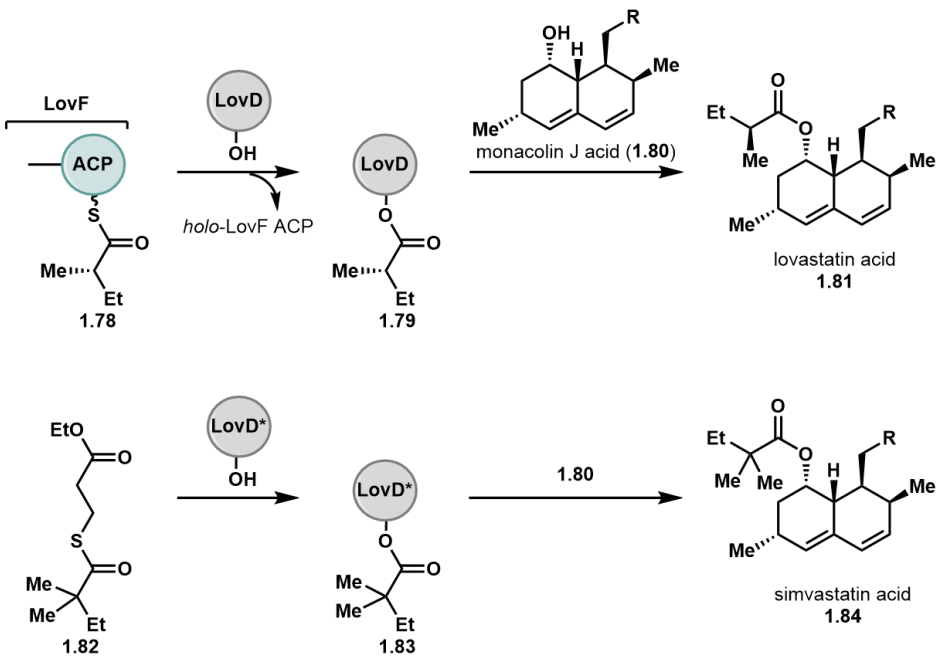


Figure 1.10. Directed evolution of the acyltransferase LovD.
Evolution from the native LovF ACP-dependent reaction (top), to increased reactivity with substrate surrogate thioester **1.82** (bottom).[146]

To the best of our knowledge, this is the only reported example of engineering toward removing a PE's CP requirement, but with the recent spotlight on directed evolution, we anticipate that more reports are forthcoming.

1.2.5. Outlook

CP-dependent biosynthetic pathways are the source of many important natural products and synthetically interesting partner enzymes. Harnessing CP-dependent PEs for biocatalysis, particularly as cascades, has the potential to help develop new routes to useful compounds. The challenges associated with PKS and NRPS PEs substrate activation are different than most other enzymes classes commonly used for biocatalysis, but the general methods presented here, fermentation, *in vitro* reconstitution, chemoenzymatic loading and directed evolution provide the basis for future efforts for wider applications of CP-dependent enzymes.

We anticipate that the most economical approaches, the chemoenzymatic methods with substrate surrogates to acylate KSs and ACPs and provide extender units for ATs will be the most fruitful in the near future. These incorporate advantages in chemical synthesis (large-scale preparation and diversification) with biosynthesis (high chemo-, site- and stereoselectivity). For PEs that are absolutely CP-dependent, building catalytic cycles to recycle CPs may be the most fruitful initially, later followed by evolution away from CPs. When our understanding of engineering PKS and NRPS modules pathways has progressed, fermentation may eventually be the most productive option.

1.3 Biocatalytic CP-PE Platforms

Although many PEs discussed in Section 1.2 have some flexibility in their native hosts and in *in vitro* reactions, bypassing native gatekeepers to use non-native substrates can be challenging. Synthetically interesting PEs are found throughout CP-dependent pathways, not just at the beginning or end. Without consideration for both loading substrates onto a CP and offloading a product, PEs are only useful for a single turnover of CP. My thesis will summarize studies with PEs to catalyze non-native reactions in the chemoenzymatic synthesis of chiral amines. My work investigates both reactions on non-native CP-bound substrates and new CP-independent pathways.

Our strategies for applying PEs synthetically must address the CP dependence, by multiple possible options: removing the dependence, exploiting a CP-independent mode, discovering a robust small molecule surrogate, or by incorporating loading and offloading into our catalytic cycles to also utilize CPs in catalytic amounts (Figure 1.11).

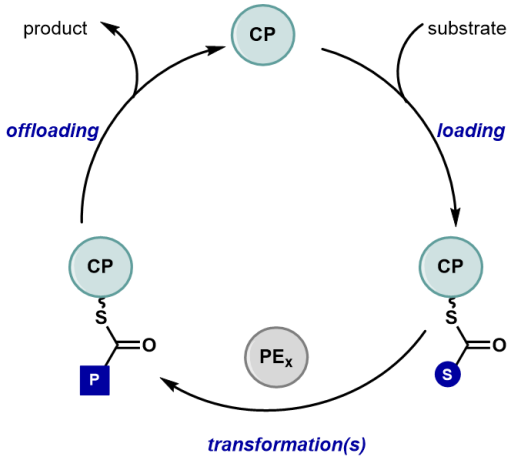


Figure 1.11. A scalable CP-PE biocatalytic platform
This will require including loading and offloading steps in the same pot such that the carrier protein is used in catalytic amounts as well.

Dissertation Outline

The following Chapters will cover my research on various CP-dependent enzymes toward developing platforms that utilize those PEs biocatalytically (Figure 1.12).

Chapter 2, "Characterization of the Polyketide-Like synthase SxtA", summarizes our studies on characterizing the native functionality on the first protein in saxitoxin biosynthesis, SxtA. This four-domain enzyme first caught our interest because it appeared to have its own loading domain and a rare C–C bond-forming α-oxoamine synthase (AOS) domain as its offloading mechanism. Through small molecule and intact protein MS, we elucidate on how SxtA's starter unit malonyl-CoA is likely loaded by a different acyltransferase protein onto the ACP domain and elaborated by two other PE domains into propionyl-ACP. Acyl groups are offloaded by the AOS protein member, the 8-amino-7-oxononanoate synthase (AONS) domain, which performs a Claisen-like condensation with propionyl-ACP and the amino acid L-Arg. Our initial characterization of the SxtA domains was then used to develop SxtA AONS as a tool for derivatizing α-amino acids and further investigate the mechanisms of the SxtA methyltransferase (MT) and AONS domains.

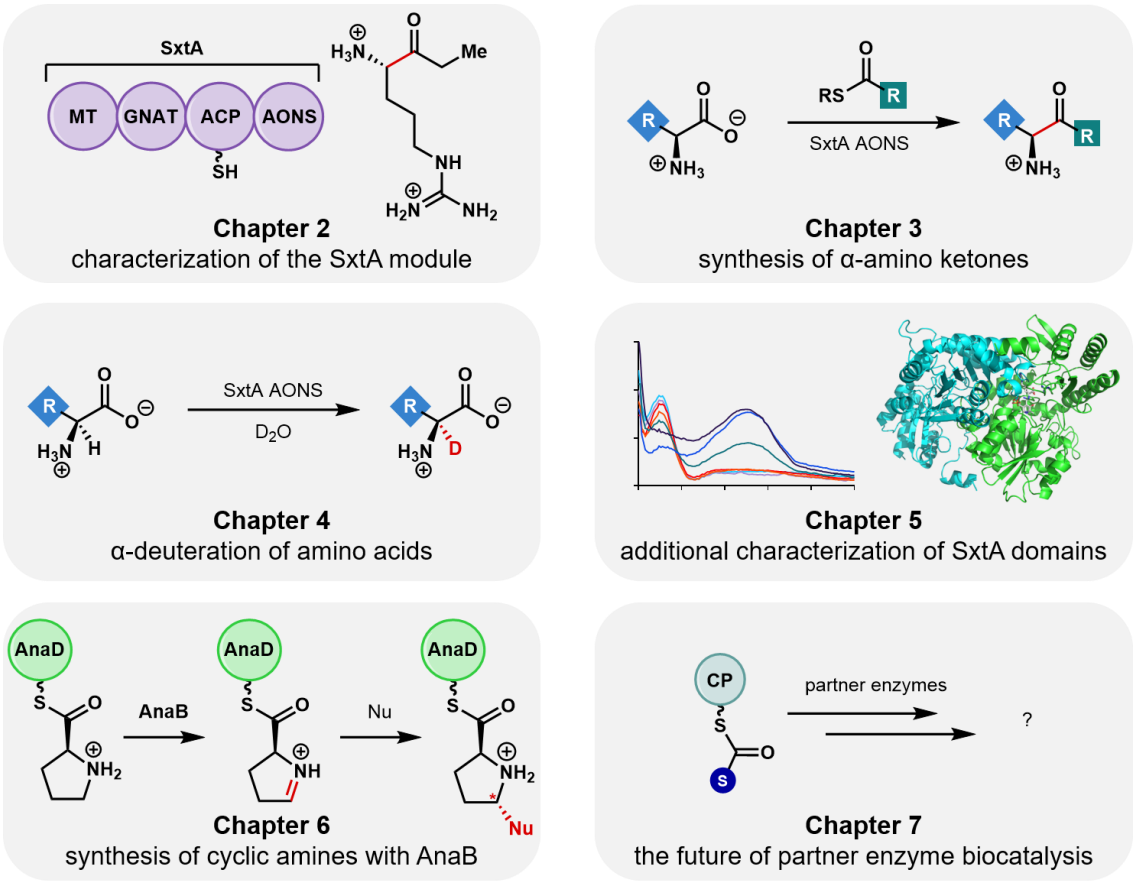


Figure 1.12. Biocatalytic reactions studied in Chapters 2-6.

Chapter 3, "Chemoenzymatic Synthesis of α-Amino Ketones", covers our efforts to utilize SxtA AONS as a biocatalyst to synthesize multiple α-amino ketones from different combinations of amino acids and acyl-thioesters. First, we optimized the conditions to generate SxtA's native arginine-based ketone product on a larger scale. Additionally, we explored SxtA AONS' ability to synthesize other ketone products from non-native amino acids and acyl group substrates. We began a directed evolution campaign to engineer a SxtA AONS variant to synthesize a tryptophan-based ketone product.

In the second direction of amino acid elaboration, we also used wild-type SxtA AONS to install deuterium at the alpha carbon of amino acids in **Chapter 4, "SxtA AONS-Mediated Deuterium**

labeling". The substrate scope for deuterium incorporation into amino acids was initially small, but esterification to methyl esters greatly increased the breadth of side chains accepted, including some D-methyl esters. A preparative-scale deuteration reaction provided deuterated alanine methyl ester for the chemoenzymatic synthesis of an isotopically labeled analog of the drug safinamide.

The last section on saxitoxin biosynthetic proteins, **Chapter 5, "Structural and Spectroscopic Characterization of SxtA"**, includes our efforts to understand the finer mechanistic details of SxtA. Two of its PE domains, MT and AONS, belong to enzyme subclasses that are underexplored. In collaboration with the laboratory of Prof. Janet Smith, we have begun exploring the structural details of SxtA MT to understand its mechanism. We also report some preliminary spectroscopic properties of SxtA AONS, which may provide the basis for improving the enzyme as a biocatalyst.

Chapter 6, "Synthesis of Chiral Cyclic Amines with Oxidase AnaB", is focused on an entirely different system. The flavin-dependent oxidase AnaB is from the type II segment of the anatoxin-a biosynthetic pathway and performs a two-electron oxidation on a prolyl substrate bound to the PCP AnaD. The resulting iminium product could potentially be attacked by a nucleophile to provide chiral cyclic amines, but without loading and offloading steps added like SxtA, the final yield would be restricted to one turnover of AnaD. Therefore, we built two catalytic cycles to utilize AnaD and AnaB catalytically, with different approaches to loading and offloading.

The concluding section, **Chapter 7, "Conclusions and Future Directions"**, will discuss my thoughts on the future of PE-mediated reactions and biocatalysis in general.

1.4 References

(1) Bornscheuer, U. T.; Huisman, G. W.; Kazlauskas, R. J.; Lutz, S.; Moore, J. C.; Robins, K. Engineering the Third Wave of Biocatalysis. *Nature* **2012**, *485*, 185–194.

(2) Sheldon, R. A.; Brady, D. The Limits to Biocatalysis: Pushing the Envelope. *Chem. Commun.* **2018**, *54*, 6088–6104.

(3) Balkenhohl, F.; Ditrich, K.; Hauer, B.; Ladner, W. Optically Active Amines via Lipase-Catalyzed Methoxyacetylation. *J. für Prakt. Chemie/Chemiker-Zeitung* **1997**, *339*, 381–384.

(4) Schmid, A.; Dordick, J. S.; Hauer, B.; Kiener, A.; Wubbolts, M.; Witholt, B. Industrial Biocatalysis Today and Tomorrow. *Nature* **2001**, *409*, 258–268.

(5) Parmeggiani, F.; Weise, N. J.; Ahmed, S. T.; Turner, N. J. Synthetic and Therapeutic Applications of Ammonia-Lyases and Aminomutases. *Chem. Rev.* **2018**, *118*, 73–118.

(6) Savile, C. K.; Janey, J. M.; Mundorff, E. C.; Moore, J. C.; Tam, S.; Jarvis, W. R.; Colbeck, J. C.; Krebber, A.; Fleitz, F. J.; Brands, J.; Devine, P. N.; Huisman, G. W.; Hughes, G. J. Biocatalytic Asymmetric Synthesis of Sitagliptin Manufacture. *Science* **2010**, *329*, 305–310.

(7) Patel, R. N. Biocatalysis for Synthesis of Pharmaceuticals. *Bioorganic Med. Chem.* **2018**, *26*, 1252–1274.

(8) Martinez, C. A.; Hu, S.; Dumond, Y.; Tao, J.; Kelleher, P.; Tully, L. Development of a Chemoenzymatic Manufacturing Process for Pregabalin. *Org. Process Res. Dev.* **2008**, *12*, 392–398.

(9) Fesko, K.; Gruber-Khadjawi, M. Biocatalytic Methods for C-C Bond Formation. *ChemCatChem* **2013**, *5*, 1248–1272.

(10) Schmidt, N. G.; Eger, E.; Kroutil, W. Building Bridges: Biocatalytic C-C-Bond Formation toward Multifunctional Products. *ACS Catal.* **2016**, *6*, 4286–4311.

(11) Nestl, B. M.; Hammer, S. C.; Nebel, B. A.; Hauer, B. New Generation of Biocatalysts for Organic Synthesis. *Angew. Chem. Int. Ed.* **2014**, *53*, 3070–3095.

(12) Hughes, R. A.; Ellington, A. D. Synthetic DNA Synthesis and Assembly: Putting the Synthetic in Synthetic Biology. *Cold Spring Harb. Perspect. Biol.* **2017**, *9*, a023812.

(13) Kosuri, S.; Church, G. M. Large-Scale de Novo DNA Synthesis: Technologies and Applications. *Nat. Methods* **2014**, *11*, 499–507.

(14) Mardis, E. R. A Decade's Perspective on DNA Sequencing Technology. *Nature* **2011**, *470*, 198–203.

(15) Mardis, E. R. The Impact of Next-Generation Sequencing Technology on Genetics. *Trends Genet.* **2008**, *24*, 133–141.

(16) Agresti, J. J.; Antipov, E.; Abate, A. R.; Ahn, K.; Rowat, A. C.; Baret, J.-C.; Marquez, M.; Klibanov, A. M.; Griffiths, A. D.; Weitz, D. A. Ultrahigh-Throughput Screening in Drop-Based Microfluidics for Directed Evolution. *Proc. Natl. Acad. Sci.* **2010**, *107*, 4004–4009.

(17) Liu, B.; Li, S.; Hu, J. Technological Advances in High-Throughput Screening. *Am. J. PharmacoGenomics* **2004**, *4*, 263–276.

(18) Arnold, F. H. Combinatorial and Computational Challenges for Biocatalyst Design. *Nature* **2001**, *409*, 253–257.

(19) Voigt, C. A.; Mayo, S. L.; Arnold, F. H.; Wang, Z.-G. Computational Method to Reduce the Search Space for Directed Protein Evolution. *Proc. Natl. Acad. Sci.* **2001**, *98*, 3778–3783.

(20) Martínez, R.; Schwaneberg, U. A Roadmap to Directed Enzyme Evolution and Screening

Systems for Biotechnological Applications. *Biol. Res.* **2013**, *46*, 395–405.
(21) Larsen, P. O.; von Ins, M. The Rate of Growth in Scientific Publication and the Decline in Coverage Provided by Science Citation Index. *Scientometrics* **2010**, *84*, 575–603.
(22) Goel, R. K.; Faria, J. R. PROLIFERATION OF ACADEMIC JOURNALS: EFFECTS ON RESEARCH QUANTITY AND QUALITY. *Metroeconomica* **2007**, *58*, 536–549.
(23) Lau, J.; Tran, C.; Licari, P.; Galazzo, J. Development of a High Cell-Density Fed-Batch Bioprocess for the Heterologous Production of 6-Deoxyerythronolide B in Escherichia Coli. *J. Biotechnol.* **2004**, *110*, 95–103.
(24) Jia, Z.; Zhang, X.; Zhao, Y.; Cao, X. Enhancement of Lovastatin Production by Supplementing Polyketide Antibiotics to the Submerged Culture of Aspergillus Terreus. *Appl. Biochem. Biotechnol.* **2010**, *160*, 2014–2025.
(25) Keatinge-Clay, A. T. The Uncommon Enzymology of Cis-Acyltransferase Assembly Lines. *Chem. Rev.* **2017**, *117*, 5334–5366.
(26) Helfrich, E. J. N.; Piel, J. Biosynthesis of Polyketides by Trans-AT Polyketide Synthases. *Nat. Prod. Rep.* **2016**, *33*, 231–316.
(27) Challis, G. L.; Naismith, J. H. Structural Aspects of Non-Ribosomal Peptide Biosynthesis. *Curr. Opin. Struct. Biol.* **2004**, *14*, 748–756.
(28) Robbins, T.; Liu, Y.-C.; Cane, D. E.; Khosla, C. Structure and Mechanism of Assembly Line Polyketide Synthases. *Curr. Opin. Struct. Biol.* **2016**, *41*, 10–18.
(29) Beld, J.; Lee, D. J.; Burkart, M. D. Fatty Acid Biosynthesis Revisited: Structure Elucidation and Metabolic Engineering. *Mol. Biosyst.* **2015**, *11*, 38–59.
(30) Shimizu, Y.; Ogata, H.; Goto, S. Type III Polyketide Synthases: Functional Classification and Phylogenomics. *Chemb* **2017**, *18*, 50–65.
(31) Du, L.; Lou, L. PKS and NRPS Release Mechanisms. *Nat. Prod. Rep.* **2010**, *27*, 255–278.
(32) Lohman, J. R.; Ma, M.; Osipiuk, J.; Nocek, B.; Kim, Y.; Chang, C.; Cuff, M.; Mack, J.; Bigelow, L.; Li, H.; Endres, M.; Babnigg, G.; Joachimiak, A.; Phillips, G. N.; Shen, B. Structural and Evolutionary Relationships of "AT-Less" Type I Polyketide Synthase Ketosynthases. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 12693–12698.
(33) Baltz, R. H. Molecular Engineering Approaches to Peptide, Polyketide and Other Antibiotics. *Nat. Biotechnol.* **2006**, *24*, 1533–1540.
(34) Sherman, D. H. The Lego-Ization of Polyketide Biosynthesis. *Nat. Biotechnol.* **2005**, *23*, 1083–1084.
(35) Klaus, M.; Grininger, M. Natural Product Reports Engineering Strategies for Rational Polyketide Synthase Design. *Nat. Prod. Rep.* **2018**, *35*, 1070–1081.
(36) Xu, W.; Qiao, K.; Tang, Y. Structural Analysis of Protein – Protein Interactions in Type I Polyketide Synthases. *Crit. Rev. Biochem. Mol. Biol.* **2013**, *48*, 98–122.
(37) Barajas, J. F.; Blake-Hedges, J. M.; Bailey, C. B.; Curran, S.; Keasling, J. D. Engineered Polyketides : Synergy between Protein and Host Level Engineering. *Synth. Syst. Biotechnol.* **2017**, *2*, 147–166.
(38) Süssmuth, R. D.; Mainz, A. Nonribosomal Peptide Synthesis — Principles and Prospects. *Angew. Chem. Int. Ed.* **2017**, *56*, 3770–3821.
(39) Winn, M.; Fyans, J. K.; Zhuo, Y.; Micklefield, J. Recent Advances in Engineering Nonribosomal Peptide Assembly Lines 2015. *Nat. Prod. Rep.* **2015**, *33*, 317–347.
(40) Klaus, M.; Ostrowski, M. P.; Austerjost, J.; Robbins, T.; Lowry, B.; Cane, D. E.; Khosla, C. Protein-Protein Interactions, Not Substrate Recognition, Dominate the Turnover of Chimeric Assembly Line Polyketide Synthases. *J. Biol. Chem.* **2016**, *291*, 16404–16415.

(41) Staunton, J.; Weissman, K. J. Polyketide Biosynthesis: A Millennium Review. *Nat. Prod. Rep.* **2001**, *18*, 380–416.
(42) Siskos, A. P.; Baerga-Ortiz, A.; Bali, S.; Stein, V.; Mamdani, H.; Spiteller, D.; Popovic, B.; Spencer, J. B.; Staunton, J.; Weissman, K. J.; Leadlay, P. F. Molecular Basis of Celmer's Rules: Stereochemistry of Catalysis by Isolated Ketoreductase Domains from Modular Polyketide Synthases. *Chem. Biol.* **2005**, *12*, 1145–1153.
(43) Piasecki, S. K.; Taylor, C. A.; Detelich, J. F.; Liu, J.; Zheng, J.; Komsoukaniants, A.; Siegel, D. R.; Keatinge-Clay, A. T. Employing Modular Polyketide Synthase Ketoreductases as Biocatalysts in the Preparative Chemoenzymatic Syntheses of Diketide Chiral Building Blocks. *Chem. Biol.* **2011**, *18*, 1331–1340.
(44) Dodge, G. J.; Ronnow, D.; Taylor, R. E.; Smith, J. L. Molecular Basis for Olefin Rearrangement in the Gephyronic Acid Polyketide Synthase. *ACS Chem. Biol.* **2018**, *13*, 2699–2707.
(45) Jiang, W.; Heemstra, J. R.; Forseth, R. R.; Neumann, C. S.; Manaviazar, S.; Schroeder, F. C.; Hale, K. J.; Walsh, C. T. Biosynthetic Chlorination of the Piperazate Residue in Kutzneride Biosynthesis by KthP. *Biochemistry* **2011**, *50*, 6063–6072.
(46) Lebel, H.; Marcoux, J.-F.; Molinaro, C.; Charette, A. B. Stereoselective Cyclopropanation Reactions. *Chem. Rev.* **2003**, *103*, 977–1050.
(47) Neumann, C. S.; Walsh, C. T. Biosynthesis of (-)-(1S,2R) -Allocoronamic Acyl Thioester by an Fe(II)-Dependent Halogenase and a Cyclopropane-Forming Flavoprotein. *J. Am. Chem. Soc.* **2008**, *130*, 14022–14023.
(48) Winter, J. M.; Behnken, S.; Hertweck, C. Genomics-Inspired Discovery of Natural Products. *Curr. Opin. Chem. Biol.* **2011**, *15*, 22–31.
(49) Fischbach, M. A.; Walsh, C. T. Assembly-Line Enzymology for Polyketide and Nonribosomal Peptide Antibiotics : Logic , Machinery , and Mechanisms. *Chem. Rev.* **2006**, *5*, 3468–3496.
(50) Fontana, A.; Manzo, E.; Ciavatta, M. L.; Gavagnin, M.; Cimino, G. *Handbook of Marine Natural Products*; Fattorusso, E., Gerwick, W. H., Taglialatela-Scafati, O., Eds.; Springer Netherlands: Dordrecht, 2012.
(51) Minas, W. Production of Erythromycin with Saccharopolyspora Erythraea. *Methods Biotechnol.* **2005**, *18*, 65–90.
(52) Chopra, I.; Roberts, M. Tetracycline Antibiotics: Mode of Action, Applications, Molecular Biology, and Epidemiology of Bacterial Resistance. *Microbiol. Mol. Biol. Rev.* **2001**, *65*, 232–260.
(53) Parekh, S. Improvement of Microbial Strains and Fermentation Processes. *Appl. Environ. Microbiol.* **2000**, 287–301.
(54) Zhang, H.; Wang, Y.; Wu, J.; Skalina, K.; Pfeifer, B. A. Complete Biosynthesis of Erythromycin A and Designed Analogs Using E. Coli as a Heterologous Host. *Chem. Biol.* **2010**, *17*, 1232–1240.
(55) MacFadyen, A. C. Determination of the Genetic Basis of Oxytetracycline Productivity for Streptomyces Rimosus , by the Examination of Genomic Data ., University of Strathclyde, 2015.
(56) Jung, H.; Kim, S.; Moon, H. Optimization of Culture Conditions and Scale-up to Pilot and Plant Scales for Vancomycin Production by Amycolatopsis Orientalis. *Biotechnol. Prod. Process Eng.* **2007**, *77*, 789–795.
(57) Cai, W.; Zhang, W. Engineering Modular Polyketide Synthases for Production of Biofuels

and Industrial Chemicals. *Curr. Opin. Biotechnol.* **2018**, *50*, 32–38.
(58) Li, C.; Hazzard, C.; Florova, G.; Ã, K. A. R. High Titer Production of Tetracenomycins by Heterologous Expression of the Pathway in a Streptomyces Cinnamonensis Industrial Monensin Producer Strain. *Metab. Eng.* **2009**, *11*, 319–327.
(59) Liu, J.; Chen, Y.; Wang, W.; Ren, M.; Wu, P.; Wang, Y.; Li, C. Engineering of an Lrp Family Regulator SACE _ Lrp Improves Erythromycin Production in Saccharopolyspora Erythraea. *Metab. Eng.* **2017**, *39*, 29–37.
(60) Rodriguez, E.; Menzella, H. G.; Gramajo, H. *Heterologous Production of Polyketides in Bacteria*, 1st ed.; Elsevier Inc., 2009; Vol. 459.
(61) Pfeifer, B. A.; Khosla, C. Biosynthesis of Polyketides in Heterologous Hosts. *Microbiol. Mol. Biol. Rev.* **2001**, *65*, 106–118.
(62) Pfeifer, B. A.; Admiraal, S. J.; Gramajo, H.; Cane, D. E.; Khosla, C. Biosynthesis of Complex Polyketides in a Metabolically Engineered Stain of E. Coli. *Science* **2001**, *291*, 1790–1792.
(63) Pfeifer, B. A.; Hu, Z.; Licari, P.; Khosla, C. Process and Metabolic Strategies for Improved Production of Escherichia Coli -Derived 6-Deoxyerythronolide B. *Appl. Environ. Microbiol.* **2002**, *68*, 3287–3292.
(64) Fang, L.; Guell, M.; Church, G. M.; Pfeifer, B. A. Heterologous Erythromycin Production Across Strain and Plasmid Construction. *Biotechnol. Progresses* **2018**, *34*, 271–276.
(65) Choi, D.; Choi, O. Y.; Shin, H. J.; Chung, D. O.; Shin, D. Y. Tylosin Production by Streptomyces Fradiae Using Raw Cornmeal in Airlift Bioreactor. *J. Microbiol. Biotechnol.* **2007**, *17*, 1071–1078.
(66) Jami, M.-S.; Barreiro, C.; García-Estrada, C.; Martín, J.-F. Proteome Analysis of the Penicillin Producer Penicillium Chrysogenum. *Mol. Cell. Proteomics* **2010**, *9*, 1182–1198.
(67) Srivastava, P.; Mishra, P.; Kundu, S. Process Strategies for Cephalosporin C Fermentation. *J. Sci. Ind. Res.* **2006**, *65*, 599–602.
(68) Barber, M. S.; Giesecke, U.; Reichert, A.; Minas, W. Industrial Enzymatic Production of Cephalosporin-Based b -Lactams Metric Ton. In *Molecular Biotechnolgy of Fungal beta-Lactam Antibiotics and Related Peptide Synthetases*; Brakhage, A. A., Ed.; Springer, Berlin, Heidelberg: Berlin, 2004; pp 179–215.
(69) Takashima, Y.; Nakajima, H.; Sonomoto, K.; Tanaka, A. Production of Daunorubicin by Immobilized Growing Streptomyces Peucetius Cells. *Appl. Microbiol. Biotechnol.* **1987**, *27*, 106–109.
(70) Villax, I. Fermentative Production of Tetracycline. 3,432,394, 1969.
(71) Zhu, T.; Cheng, X.; Liu, Y.; Deng, Z.; You, D. Deciphering and Engineering of the Final Step Halogenase for Improved Chlortetracycline Biosynthesis in Industrial Streptomyces Aureofaciens. *Metab. Eng.* **2013**, *19*, 69–78.
(72) Gatenbeck, S.; Mosbach, K. Acetate Carboxyl Oxygen (18O) as Donor for Phenolic Hydroxy Groups of Orsellinic Acid Produced by Fungi. *Acta Chem. Scand.* **1959**, *13*, 1561–1564.
(73) Birch, A. J.; Djerassi, C.; Dutcher, J. D.; Majer, J.; Perlman, D.; Pride, E.; Rickards, R. W. Studies in Relation to Biosynthesis. Part XXXV. Macrolide Antibiotics. Part XII. Methymycin. *J. Chem. Soc.* **1964**, *0*, 5274–5278.
(74) Edwards, D. J.; Marquez, B. L.; Nogle, L. M.; McPhail, K.; Goeger, D. E.; Roberts, M. A.; Gerwick, W. H. Structure and Biosynthesis of the Jamaicamides, New Mixed Polyketide-Peptide Neurotoxins from the Marine Cyanobacterium Lyngbya Majuscula. *Chem. Biol.*

2014, *128*, 189–190.
(75) Bode, H. B.; Brachmann, A. O.; Jadhav, K. B.; Seyfarth, L.; Dauth, C.; Fuchs, S. W.; Kaiser, M.; Waterfield, N. R.; Sack, H.; Heinemann, S. H.; Arndt, H.-D. Structure Elucidation and Activity of Kolossin A , the D-/L-Pentadecapeptide Product of a Giant Nonribosomal Peptide Synthetase. *Angew. Chem. Int. Ed.* **2015**, *54*, 10352–10355.
(76) Houck, D. R.; Ondeyka, J.; Zink, D. L.; Inamine, E.; Goetz, M. A.; Hensens, O. D. On the Biosynthesis of Asperlicin and the Directed Biosynthesis of Analogs in Aspergillus Alliaceus. *J. Antibiot. (Tokyo).* **1988**, No. July, 882–891.
(77) Traber, R.; Hofmannand, H.; Kobel, H. Cyclosporins - New Analogues by Precursor-Directed Biosynthesis. *J. Antibiot.* **1988**, *XLII*, 591–597.
(78) Weist, S.; Bister, B.; Puk, O.; Bischoff, D.; Pelzer, S.; Nicholson, G. J.; Wohlleben, W.; Jung, G.; Süssmuth, R. D. Fluorobalhimycin – A New Chapter in Glycopeptide Antibiotic Research. *Angew. Chem. Int. Ed.* **2002**, *41*, 3383–3385.
(79) Nett, M.; Gulder, T. A. M.; Kale, A. J.; Hughes, C. C.; Moore, B. S. Function-Oriented Biosynthesis of β -Lactone Proteasome Inhibitors in Salinispora Tropica. *J. Med. Chem.* **2009**, *52*, 6163–6167.
(80) Skiba, M. A.; Maloney, F. P.; Dan, Q.; Fraley, A. E.; Aldrich, C. C.; Smith, J. L.; Brown, W. C. *PKS–NRPS Enzymology and Structural Biology: Considerations in Protein Production*, 1st ed.; Elsevier Inc., 2018; Vol. 604.
(81) Rodionov, D. A.; Kurnasov, O. V.; Stec, B.; Wang, Y.; Roberts, M. F.; Osterman, A. L. Genomic Identification and in Vitro Reconstitution of a Complete Biosynthetic Pathway for the Osmolyte Di-Myo-Inositol-Phosphate. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 4279–4284.
(82) Hansen, D. A.; Rath, C. M.; Eisman, E. B.; Narayan, A. R. H.; Kittendorf, J. D.; Mortison, J. D.; Yoon, Y. J.; Sherman, D. H. Biocatalytic Synthesis of Pikromycin, Methymycin, Neomethymycin, Novamethymycin, and Ketomethymycin. *J. Am. Chem. Soc.* **2013**, *135*, 11232–11238.
(83) Sundlov, J. A.; Shi, C.; Wilson, D. J.; Aldrich, C. C.; Gulick, A. M. Structural and Functional Investigation of the Intermolecular Interaction between NRPS Adenylation and Carrier Protein Domains. *Chem. Biol.* **2012**, *19*, 188–198.
(84) Thomas, M. G.; Burkart, M. D.; Walsh, C. T. Conversion of L-Proline to Pyrrolyl-2-Carboxyl-S-PCP during Undecylprodigiosin and Pyoluteorin Biosynthesis. *Chem. Biol.* **2002**, *9*, 171–184.
(85) Dorrestein, P. C.; Yeh, E.; Garneau-Tsodikova, S.; Kelleher, N. L.; Walsh, C. T. Dichlorination of a Pyrrolyl-S-Carrier Protein by FADH2-Dependent Halogenase PltA during Pyoluteorin Biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 13843–13848.
(86) Méjean, A.; Mann, S.; Vassiliadis, G.; Lombard, B.; Loew, D.; Ploux, O. In Vitro Reconstitution of the First Steps of Anatoxin-a Biosynthesis in Oscillatoria PCC 6506: From Free L-Proline to Acyl Carrier Protein Bound Dehydroproline. *Biochemistry* **2010**, *49*, 103–113.
(87) Mann, S.; Lombard, B.; Loew, D.; Méjean, A.; Ploux, O. Insights into the Reaction Mechanism of the Prolyl - Acyl Carrier Protein Oxidase Involved in Anatoxin-a and Homoanatoxin-a Biosynthesis. *Biochemistry* **2011**, *50*, 7184–7197.
(88) Stachelhaus, T.; Walsh, C. T. Mutational Analysis of the Epimerization Domain in the Initiation Module PheATE. *Biochemistry* **2000**, *39*, 5775–5787.
(89) Greunke, C.; Glçckle, A.; Antosch, J.; Gulder, T. A. M. Natural Product Synthesis

Biocatalytic Total Synthesis of Ikarugamycin. *Angew. Chem. Int. Ed.* **2017**, *56*, 4351–4355.
(90) Shen, B.; Hutchinson, C. R.; Shen, B.; Hutchinson, C. R. Enzymatic Synthesis of a Bacterial Polyketide from Acetyl and Malonyl Coenzyme A. *Science* **1993**, *262*, 1535–1540.
(91) Carreras, C. W.; Pieper, R.; Khosla, C. Efficient Synthesis of Aromatic Polyketides in Vitro by the Actinorhodin Polyketide Synthase. *J. Am. Chem. Soc.* **1996**, *7863*, 5158–5159.
(92) Carreras, C. W.; Khosla, C. Purification and in Vitro Reconstitution of the Essential Protein Components of an Aromatic Polyketide Synthase. *Biochemistry* **1998**, *2960*, 2084–2088.
(93) Sun, Y.; Hahn, F.; Demydchuk, Y.; Chettle, J.; Tosin, M.; Osada, H.; Leadlay, P. F. In Vitro Reconstruction of Tetronate RK-682 Biosynthesis. *Nat. Chem. Biol.* **2009**, *6*, 99–101.
(94) Cochrane, R. V. K.; Sanichar, R.; Lambkin, G. R.; Xu, W.; Vederas, J. C.; Tang, Y. Identification and Reconstitution of the Polyketide Synthases Responsible for Biosynthesis of the Anti-Malarial Agent, Cladosporin. *Angew. Chem. Int. Ed.* **2016**, *55*, 664–668.
(95) Lowry, B.; Robbins, T.; Weng, C.-H.; Brien, R. V. O.; Cane, D. E. In Vitro Reconstitution and Analysis of the 6-Deoxyerythronolide B Synthase. *J. Am. Chem. Soc.* **2013**, *135*, 16809–16812.
(96) Dan, Q.; Newmister, S. A.; Klas, K. R.; Fraley, A. E.; McAfoos, T. J.; Somoza, A. D.; Sunderhaus, J. D.; Ye, Y.; Shende, V. V.; Yu, F.; Sanders, J. N.; Brown, W. C.; Zhao, L.; Paton, R. S.; Houk, K. N.; Smith, J. L.; Sherman, D. H.; Williams, R. M. Fungal Indole Alkaloid Biogenesis through Evolution of a Bifunctional Reductase/Diels–Alderase. *Nat. Chem.* **2019**, *11*, 972–980.
(97) Choi, M. S.; Houk, K. N.; Tang, Y.; Gao, X.; Jiang, W.; Jime, G.; Walsh, C. T. An Iterative, Bimodular Nonribosomal Peptide Synthetase That Converts Anthranilate and Tryptophan into Tetracyclic Asperlicins. *Chem. Biol.* **2013**, *20*, 870–878.
(98) Keating, T. A.; Marshall, C. G.; Walsh, C. T. Reconstitution and Characterization of the Vibrio Cholerae Vibriobactin Synthetase from VibB, VibE, VibF, and VibH. *Biochemistry* **2000**, *39*, 15522–15530.
(99) Paquette, L. A.; Romine, J. L.; Lin, H.-S.; Wright, J. Total Synthesis of (+)-Ikarugamycin. 1. Stereocontrolled Construction of the Decahydro-as-Indacene Subunit. *J. Am. Chem. Soc.* **1990**, *112*, 9284–9292.
(100) Paquette, L. A.; Macdonald, D.; Anderson, L. G. Total Synthesis of (+)-Ikarugamycin Elaboration of the Macrocyclic Lactam and Tetramic Acid Substructures and Complete Assembly of the Antibiotic. *J. Am. Chem. Soc.* **1990**, No. 7, 9292–9299.
(101) Cheng, Q.; Xiang, L.; Izumikawa, M.; Meluzzi, D.; Moore, B. S. Enzymatic Total Synthesis of Enterocin Polyketides. *Nat. Chem. Biol.* **2007**, *3*, 557–558.
(102) Rizzo, A.; Trauner, D. Toward (−)-Enterocin: An Improved Cuprate Barbier Protocol To Overcome Strain and Sterical Hindrance. *Org. Lett.* **2018**, *20*, 1841–1844.
(103) Sattely, E. S.; Fischbach, M. A.; Walsh, C. T. Total Biosynthesis: In Vitro Reconstitution of Polyketide and Nonribosomal Peptide Pathways. *Nat. Prod. Rep.* **2008**, *25*, 757.
(104) Lowry, B.; Walsh, C. T.; Khosla, C. In Vitro Reconstitution of Metabolic Pathways: Insights into Nature's Chemical Logic. *Synlett* **2015**, *26*, 1008–1025.
(105) Franke, J.; Hertweck, C. Biomimetic Thioesters as Probes for Enzymatic Assembly Lines: Synthesis, Applications, and Challenges. *Cell Chem. Biol.* **2016**, *23*, 1179–1192.
(106) Manandhar, M.; Cronan, J. E. A Canonical Biotin Synthesis Enzyme, 8-Amino-7-Oxononanoate Synthase (BioF), Utilizes Different Acyl Chain Donors in Bacillus Subtilis and Eschericia Coli. *Appl. Environ. Microbiol.* **2018**, *84*, e02084-17.
(107) Roberts, D. M.; Bartel, C.; Scott, A.; Ivison, D.; Simpson, T. J.; Cox, R. J. Catalytic Domain

from an Iterative Highly Reducing Fungal Polyketide Synthase Revealse Key Components of Programming. *Chem. Sci.* **2017**, *8*, 1116–1126.
(108) Kass, L. R.; Brock, D. J. H.; Bloch, K. Beta-Hydroxydecanoyl Thioester Dehydrase: I. Purification and Properties. *J. Biol. Chem.* **1967**, *242*, 4418–4432.
(109) Stevens, D. C.; Wagner, D. T.; Manion, H. R.; Alexander, B. K.; Keatinge-Clay, A. T. Methyltransferases Excised from Trans-AT Polyketide Synthases Operate on N-Acetylcysteamine-Bound Substrates. *J. Antibiot. (Tokyo).* **2016**, *69*, 1–4.
(110) Ding, Y.; Rath, C. M.; Bolduc, K. L.; Håkansson, K.; Sherman, D. H. Chemoenzymatic Synthesis of Cryptophycin Anticancer Agents by an Ester Bond-Forming Non-Ribosomal Peptide Synthetase Module. *J. Am. Chem. Soc.* **2011**, *133*, 14492–14495.
(111) Berkhan, G.; Hahn, F. A Dehydratase Domain in Ambruticin Biosynthesis Displays Additional Activity as a Pyran-Forming Cyclase. *Angew. Chem. Int. Ed.* **2014**, *53*, 14240–14244.
(112) Schneider, T. L.; Shen, B.; Walsh, C. T. Oxidase Domains in Epothilone and Bleomycin Biosynthesis : Thiazoline to Thiazole Oxidation during Chain Elongation. *Biochemistry* **2003**, 9722–9730.
(113) Sims, J. W.; Schmidt, E. W. Thioesterase-Like Role for Fungal PKS-NRPS Hybrid Reductive Domains. *J. Am. Chem. Soc.* **2008**, *130*, 11149–11155.
(114) Cheng, Y.-Q.; Tang, G.-L.; Shen, B. Type I Polyketide Synthase Requiring a Discrete Acyltransferase for Polyketide Biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 3149–3154.
(115) Riva, E.; Wilkening, I.; Gazzola, S.; Li, W. M. A.; Smith, L.; Leadlay, P. F.; Tosin, M. Chemical Probes for the Functionalization of Polyketide Intermediates. *Angew. Chem. Int. Ed.* **2014**, *53*, 11944–11949.
(116) Ad, O.; Thuronyi, B. W.; Chang, M. C. Y. Elucidating the Mechanism of Fluorinated Extender Unit Loading for Improved Production of Fluorine-Containing Polyketides. *Proc. Natl. Acad. Sci.* **2017**, *114*, E660–E668.
(117) Lowell, A. N.; DeMars, M. D.; Slocum, S. T.; Yu, F.; Anand, K.; Chemler, J. A.; Korakavi, N.; Priessnitz, J. K.; Park, S. R.; Koch, A. A.; Schultz, P. J.; Sherman, D. H. Chemoenzymatic Total Synthesis and Structural Diversification of Tylactone-Based Macrolide Antibiotics through Late-Stage Polyketide Assembly, Tailoring, and C-H Functionalization. *J. Am. Chem. Soc.* **2017**, *139*, 7913–7920.
(118) Pohl, N. L.; Gokhale, R. S.; Cane, D. E.; Khosla, C. Synthesis and Incorporation of an N-Acetylcysteamine Analogue of Methylmalonyl-CoA by a Modular Polyketide Synthase. *J. Am. Chem. Soc.* **1998**, *120*, 11206–11207.
(119) Dunbar, K. L.; Scharf, D. H.; Litomska, A.; Hertweck, C. Enzymatic Carbon-Sulfur Bond Formation in Natural Product Biosynthesis. *Chem. Rev.* **2017**, *117*, 5521–5577.
(120) Sieber, S. A.; Tao, J.; Walsh, C. T.; Marahiel, M. A. Peptidyl Thiophenols as Substrates for Nonribosomal Peptide Cyclases. *Angew. Chem. Int. Ed.* **2004**, *43*, 493–498.
(121) Haynes, S. W.; Ames, B. D.; Gao, X.; Tang, Y.; Walsh, C. T. Unraveling Terminal C-Domain-Mediated Condensation in Fungal Biosynthesis of Imidazoindolone Metabolites. *Biochemistry* **2011**, *50*, 5668–5679.
(122) Foulke-Abel, J.; Townsend, C. A. Demonstration of Starter Unit Interprotein Transfer from a Fatty Acid Synthase to a Multidomain , Nonreducing Polyketide Synthase. *ChemBioChem* **2012**, 1880–1884.
(123) Hiratsuka, T.; Suzuki, H.; Kariya, R.; Seo, T.; Minami, A.; Oikawa, H. Biosynthesis of the

Structurally Unique Polycyclopropanated Polyketide-Nucleoside Hybrid Jawsamycin (FR-900848). *Angew. Chem. Int. Ed.* **2014**, *53*, 5423–5426.
(124) Hansen, D. A.; Koch, A. A.; Sherman, D. H. Substrate Controlled Divergence in Polyketide Synthase Catalysis. *J. Am. Chem. Soc.* **2015**, *137*, 3735–3738.
(125) Meier, J. L.; Burkart, M. D. The Chemical Biology of Modular Biosynthetic Enzymes. *Chem. Soc. Rev.* **2009**, *38*, 2012.
(126) Sánchez, C.; Du, L.; Edwards, D. J.; Toney, M. D.; Shen, B. Cloning and Characterization of a Phosphopantetheinyl Transferase from Streptomyces Verticillus ATCC15003, the Producer of the Hybrid Peptide–Polyketide Antitumor Drug Bleomycin. *Chem. Biol.* **2001**, *8*, 725–738.
(127) Quadri, L. E. N.; Weinreb, P. H.; Lei, M.; Nakano, M. M.; Zuber, P.; Walsh, C. T. Characterization of Sfp, a Bacillus Subtilis Phosphopantetheinyl Transferase for Peptidyl Carrier Protein Domains in Peptide Synthetases. *Biochemistry* **1998**, *37*, 1585–1595.
(128) Beld, J.; Sonnenschein, E. C.; Vickery, C. R.; Noel, J. P.; Burkart, M. D. The Phosphopantetheinyl Transferases: Catalysis of a Post-Translational Modification Crucial for Life. *Nat. Prod. Rep.* **2014**, *31*, 61–108.
(129) Amoroso, J. W.; Borketey, L. S.; Prasad, G.; Schnarr, N. A. Direct Acylation of Carrier Proteins with Functionalized Beta-Lactones. *Org. Lett.* **2010**, *12*, 2008–2011.
(130) Prasad, G.; Amoroso, J. W.; Borketey, L. S.; Schnarr, N. A. N-Activated β -Lactams as Versatile Reagents for Acyl Carrier Protein Labeling. *Org. Biomol. Chem.* **2012**, *10*, 1992–2002.
(131) Chun, S. W.; Hinze, M. E.; Skiba, M. A.; Narayan, A. R. H. Chemistry of a Unique Polyketide-like Synthase. *J. Am. Chem. Soc.* **2018**, *140*, 2430–2433.
(132) Chen, H.; O'Connor, S.; Cane, D. E.; Walsh, C. T. Epothilone Biosynthesis: Assembly of the Methylthiazolylcarboxy Starter Unit on the EpoB Subunit. *Chem. Biol.* **2001**, *8*, 899–912.
(133) Katsuyama, Y.; Horinouchi, S. Microbial Type III Polyketide Synthases. In *Comprehensive Natural Products II*; Elsevier, 2010; Vol. 1, pp 147–170.
(134) Whicher, J. R.; Dutta, S.; Hansen, D. A.; Hale, W. A.; Chemler, J. A.; Dosey, A. M.; Narayan, A. R. H.; Håkansson, K.; Sherman, D. H.; Smith, J. L.; Skiniotis, G. Structural Rearrangements of a Polyketide Synthase Module during Its Catalytic Cycle. *Nature* **2014**, *510*, 560–564.
(135) Dutta, S.; Whicher, J. R.; Hansen, D. A.; Hale, W. A.; Chemler, J. A.; Congdon, G. R.; Narayan, A. R. H.; Håkansson, K.; Sherman, D. H.; Smith, J. L.; Skiniotis, G. Structure of a Modular Polyketide Synthase. *Nature* **2014**, *510*, 512–517.
(136) Sieber, S. a; Marahiel, M. a. Learning from Nature's Drug Factories: Nonribosomal Synthesis of Macrocyclic Peptides. *J. Bacteriol.* **2003**, *185*, 7036–7043.
(137) Kosa, N. M.; Haushalter, R. W.; Smith, A. R.; Burkart, M. D. Reversible Labeling of Native and Fusion-Protein Motifs. *Nat. Methods* **2012**, *9*, 981–984.
(138) Gulick, A. M.; Aldrich, C. C. Trapping Interactions between Catalytic Domains and Carrier Proteins of Modular Biosynthetic Enzymes with Chemical Probes. *Nat. Prod. Rep.* **2018**, *35*, 1156–1184.
(139) Sieber, S. A.; Walsh, C. T.; Marahiel, M. A. Loading Peptidyl-Coenzyme A onto Peptidyl Carrier Proteins: A Novel Approach in Characterizing Macrocyclization by Thioesterase Domains. *J. Am. Chem. Soc.* **2003**, *125*, 10862–10866.
(140) Worthington, A. S.; Burkart, M. D. One-Pot Chemo-Enzymatic Synthesis of Reporter-

Modified Proteins. *Org. Biomol. Chem.* **2006**, *4*, 44–46.
(141) Cronan, J. E.; Klages, A. L. Chemical Synthesis of Acyl Thioesters of Acyl Carrier Protein with Native Structure. *Proc. Natl. Acad. Sci.* **1981**, *78*, 5440–5444.
(142) Crosby, J.; Byrom, K. J.; Hitchman, T. S.; Cox, R. J.; Crump, M. P.; Findlow, I. S. C.; Bibb, M. J.; Simpson, T. J. Acylation of Streptomyces Type II Polyketide Synthase Acyl Carrier Proteins. *FEBS Lett.* **1998**, *433*, 132–138.
(143) Lin, T.-Y.; Borketey, L. S.; Prasad, G.; Waters, S. A.; Schnarr, N. A. Sequence, Cloning, and Analysis of the Fluvirucin B1 Polyketide Synthase from Actinomadura Vulgaris. *ACS Synth. Biol.* **2013**, *2*, 635–642.
(144) Xie, X.; Watanabe, K.; Wojcicki, W. A.; Wang, C. C. C.; Tang, Y. Biosynthesis of Lovastatin Analogs with a Broadly Specific Acyltransferase. *Chem. Biol.* **2006**, *13*, 1161–1169.
(145) Xie, X.; Tang, Y. Efficient Synthesis of Simvastatin by Use of Whole-Cell Biocatalysis. *Appl. Environ. Microbiol.* **2007**, *73*, 2054–2060.
(146) Gao, X.; Xie, X.; Pashkov, I.; Sawaya, M. R.; Laidman, J.; Zhang, W.; Cacho, R.; Yeates, T. O.; Tang, Y. Directed Evolution and Structural Characterization of a Simvastatin Synthase. *Chem. Biol.* **2009**, *16*, 1064–1074.

Chapter 2: Characterization of the Polyketide-Like Synthase SxtA

Summary

As in the case of many complex natural products, the intricate architecture of saxitoxin (STX) has hindered full exploration of this scaffold's utility as a tool for studying voltage-gated sodium ion channels and as a pharmaceutical agent. Established chemical strategies can provide access the natural product; however, a chemoenzymatic route to saxitoxin that could provide expedited access to related compounds has not been devised. The first step toward realizing a chemoenzymatic approach toward this class of molecules is the elucidation of the saxitoxin biosynthetic pathway. To date, a biochemical link between STX and its putative biosynthetic enzymes has not been demonstrated. Herein, we report the first biochemical characterization of any enzyme involved in STX biosynthesis. Specifically, the chemical functions of a polyketide-like synthase, SxtA, from the cyanobacteria *Cylindrospermopsis raciborskii* T3, are elucidated. This unique megasynthase is comprised of four domains: methyltransferase (MT), GCN5 *N*-acetyltransferase (GNAT), acyl carrier protein (ACP), and a rare example of an 8-amino-7-oxononanoate synthase (AONS) associated with a multidomain synthase. We have stablished that this single polypeptide carries

out the formation of two carbon–carbon bonds, two decarboxylation events and a stereospecific protonation to afford the linear biosynthetic precursor to STX from malonyl-coenzyme A, *S*-adenosylmethionine and L-arginine.

2.1 Introduction

SxtA, a carrier protein (CP)-containing enzyme from the biosynthetic pathway of saxitoxin (STX) is the major focus of this dissertation. STX is the causative agent of paralytic shellfish poisoning (PSP), an illness that was first documented in 1793 during George Vancouver's exploratory voyages.[1] In his 1798 book, Vancouver recounted how one crewmember passed away and four others fell ill after consuming shellfish in what is now British Columbia. Over the next 180+ years, multiple instances of PSP outbreaks were reported along the North American Pacific coasts, even as researchers continued to discover more about STX.

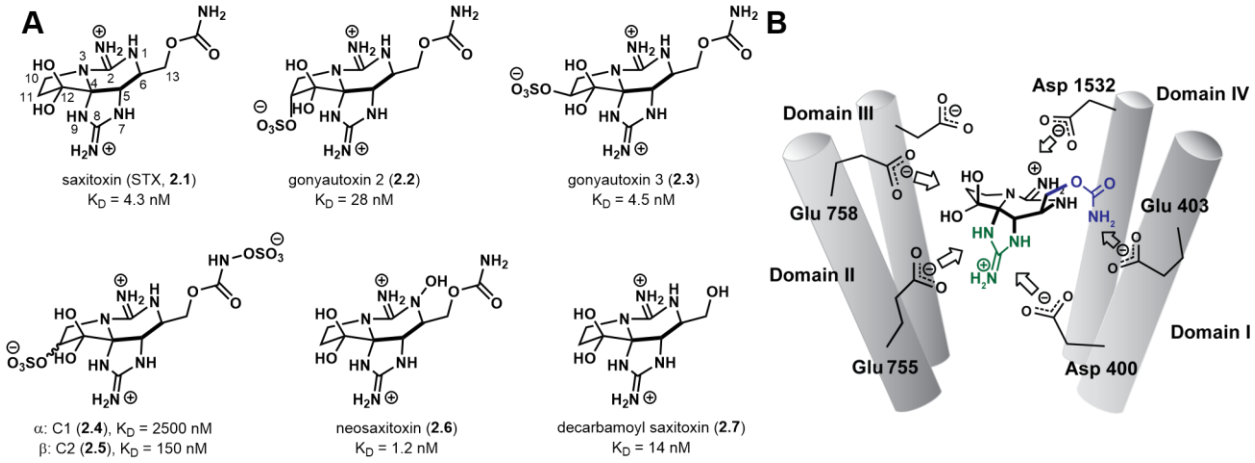


Figure 2.1. The paralytic shellfish toxin saxitoxin.
(A) The structure of saxitoxin (**2.1**) and a selection of related derivatives. IC_{50} values listed are for sodium channels in rat skeletal muscle.[2] (B) STX blocks the pore of voltage-gated sodium ion channels. Figure adapted from Shinohara *et al*.[3]

In the early to mid-1800s, scientists speculated that PSP was caused by an organic small molecule produced by mussels and clams.[4] Sommer first demonstrated in 1937 that the toxin, then often referred to as "mussel poison", was produced by some marine dinoflagellates, and that filter feeders, such as the Alaskan butter clam (*Saxidomus giganteus*), accumulated the toxin in their tissues.[5,6] Later studies found that some freshwater cyanobacteria also produced the PSP toxin.[7] The first pure sample of toxin was isolated from shellfish by Schantz in 1957, which also allowed

for the determination of the molecular formula, $C_{10}H_{17}O_4N_7 \cdot 2HCl$.[8,9] Saxitoxin was officially named in 1962, but due to difficulties with its size and amorphous salt forms, it was not until 1975 that Clardy/Schantz and Rapoport separately reported x-ray crystal structures (**2.1**, Figure 2.1A) of the molecule.[10,11] The structures revealed STX to be a tricycle with one hydrated ketone, one carbamoyl and two guanidinium groups, accounting for the compound's extremely polar nature. Kishi's racemic total synthesis of STX quickly followed in 1977.[12]

Table 2.1. Function of different human Na_v isoforms 1.1-1.9 and sensitivity to STX

Human Na_v isoform	IC_{50} of STX (nM)	Associated diseases (partial list) [7,13]
Na_v1.1	2.3	Epilepsy, migraine, autism
Na_v1.2	1.0	Epilepsy, autism
Na_v1.3	13	Epilepsy
Na_v1.4	19	Paralysis, myotonia
Na_v1.5	21	Cardiac arrythmia
Na_v1.6	1.1	Central nervous system disorders
Na_v1.7	410	Pain sensation
Na_v1.8	>10,000	Pain sensation
Na_v1.9	>10,000	Pain sensation

Subsequently, STX's mechanism of action in humans has been established as binding of voltage-gated sodium ion channels (Na_v).[7] STX binds near the outer lip of Na_v pores, where the positively charged molecule interacts strongly with the negatively charged channel residues (Figure 2.1B) to block the inward flow of Na^+ ions. This leads to respiratory paralysis and sometimes death. However, STX and its derivatives' specificity of binding to rat skeletal muscle (Figure 2.1A)[2] and different Na_v isoforms varies widely (Table 2.1).[14] Naturally occurring STX derivatives, the gonyautoxins have been investigated previously for the treatment of muscle tension (**2.2**) and pain (**2.3**), respectively, and research is ongoing to discover potentially therapeutic saxitoxins that selectively target Na_v isoforms, with particular focus on pain relief.[13,15,16]

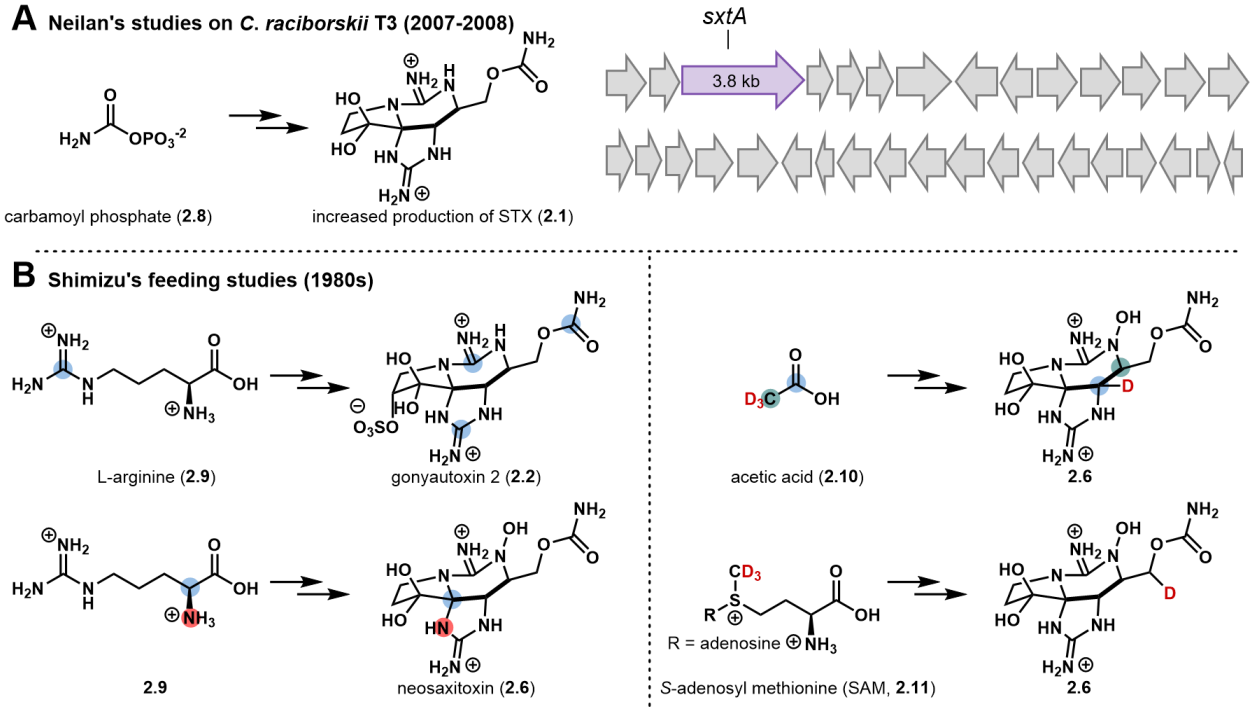


Figure 2.2. Early studies on STX biosynthesis.
(A) Neilan's studies with *C. raciborskii* T3 to discover the first *sxt* biosynthetic gene cluster.[17,18] (B) Shimizu's feeding studies with labeled precursors.[19]

Obtaining sufficient amounts of saxitoxin and analogs for assays and potential clinical trials remains a challenge. Isolation from natural sources requires a huge amount of sample (5000 L of seawater for 90 mg of impure STX, or 8 American tons of butter clams for 1 g);[1] this and *in vitro* biosynthesis with toxin-producing cyanobacterial or dinoflagellate extracts are limited to naturally occurring molecules. Asymmetric syntheses of STX are long, requiring protecting group manipulations for the numerous heteroatoms.[7] However, new derivatives could potentially be produced using biosynthetic machinery and more sustainably through manipulation of saxitoxin biosynthetic proteins. Four cyanobacterial putative gene clusters had been sequenced at the beginning of our studies in 2016 and a pathway had been proposed, though none of the enzymes had been biochemically characterized.[18,20,21]

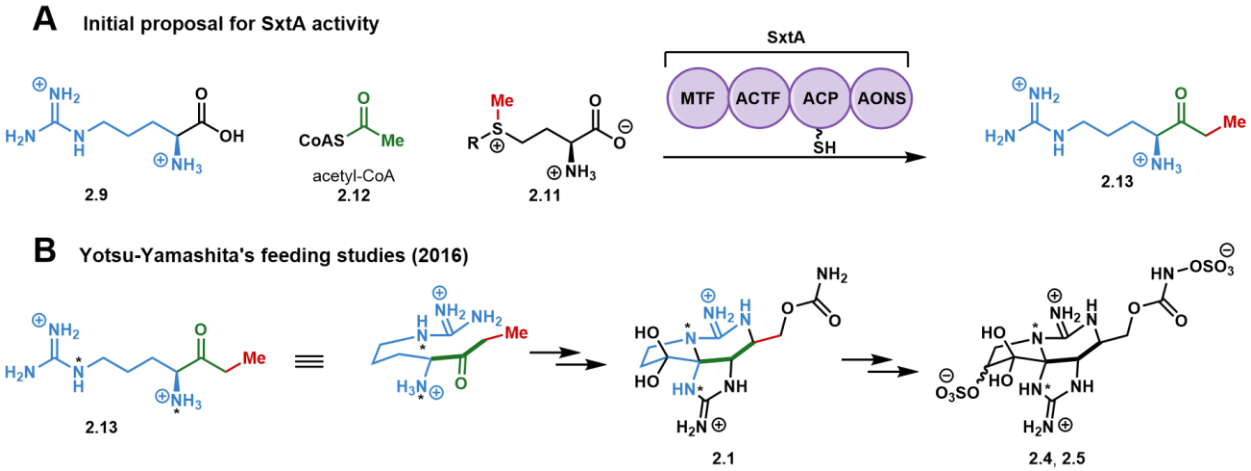


Figure 2.3. Ketone 2.13, an intermediate in STX biosynthesis.
(A) The first proposal of SxtA's substrates and native product, ketone **2.13**, by Neilan and coworkers in 2008.[18] (B) Confirmation by the Yotsu-Yamashita group of **2.13** into saxitoxins.[22] MTF: methyltransferase; ACTF: acetyltransferase; ACP: acyl carrier protein; AONS: 8-amino-7-oxononanoate synthase.

Prior to their report of the first STX biosynthetic gene cluster, Neilan and coworkers added potential STX precursors to clarified cell lysates of toxin-producing cyanobacteria *Cylindrospermopsis raciborskii* T3 and found that STX biosynthesis likely required carbamoyl phosphate (Figure 2.2A).[17] Sequencing of the *C. raciborskii* genome identified an *O*-carbamoyltransferase and a surrounding gene cluster postulated to be responsible for STX biosynthesis, which was then named *sxt*.[18] From the hypothetical protein families in the *sxt* cluster and feeding experiments by Shimizu and others (Figure 2.2B),[19] the Neilan group proposed a STX biosynthetic pathway. The largest gene, *sxtA*, encodes for a polyketide synthase-related protein with four domains predicted to initiate STX biosynthesis by producing ketone **2.13** from acetyl-CoA, *S*-adenosylmethionine (SAM), and L-arginine (Figure 2.3). Supporting evidence for **2.13** as an intermediate to STX was reported by Yotsu-Yamashita and coworkers, who identified **2.13** in lysates of toxin-producing microorganisms[23] and that an ^{15}N-labeled precursor was incorporated into labeled STX derivatives (see **2.4** and **2.5**, Figure 2.3).[22]

SxtA was the most attractive protein for initial investigations not only as the first protein in the STX biosynthetic pathway, but also as the only CP-dependent protein in the cluster. The enzyme appeared to be a single self-contained module with its own loading and offloading domains and to catalyze formation of a C–C bond stereospecifically. Additionally, SxtA's fourth domain was predicted to catalyze the transformation of an α-amino acid to α-amino ketone, a synthetically challenging process (see section 2.2 and Chapter 3). We first intended to search for the annotation each domain within the BLAST (Basic Local Alignment Search Tool) database to possibly update Neilan's proposal, then perform *in vitro* experiments with the SxtA module and finally verify the activity of each domain with individual subcloned standalone proteins.

2.2 Previously Known Details about SxtA Domain Classes

At the time of our initial studies of SxtA in 2016, four cyanobacterial homologs had been deposited in the Uniprot database. With the help of Dr. Meredith Skiba, the boundaries of SxtA's four domains were roughly defined. Performing a BLAST search on each domain turned up the following results: (1) SxtA contained no ketosynthase (KS) domain to condense multiple ketide units, classifying the protein as a polyketide-like synthase rather than a true PKS; and (2) the four SxtA domains, from the N- to C-termini, were a putative methyltransferase, a GCN5-related *N*-acetyltransferase, an acyl carrier protein, and an α-oxoamine synthase. There did not appear to be any major changes in the BLAST database from the Neilan group's findings from 2008, but we chose to adjust the domain abbreviations from MTF-ACTF-ACP-AONS to the more commonly seen MT-GNAT-ACP-AONS.

SxtA's second domain, the GCN5-related *N*-acetyltransferase (GNAT), belongs to the GNAT superfamily. Identified in all kingdoms of life, GNATs have low sequence identity to each other

across this family but typically transfer an acetyl group from acetyl-CoA to both small molecule and protein targets.[24] By sequence, SxtA GNAT is most closely related to other GNAT domains within PKS proteins. The most well-characterized member of this subclass is CurA GNAT from the curacin A biosynthetic pathway, which has been studied by three labs at the University of Michigan (Håkansson, Sherman and Smith).[25] CurA GNAT is found in the loading module of CurA, and was reported to decarboxylate malonyl-CoA (**2.14**) rapidly to acetyl-CoA (**2.15**), and load the acetyl group onto *holo*-ACP relatively slowly (Figure 2.4A). Decarboxylation of malonyl-ACP to acetyl-ACP was also noted, but the rate was slower than that of malonyl-CoA. Loading directly from acetyl-CoA without malonyl-CoA present was observed as well (i.e., **2.15** to **2.16** in Figure 2.4A). Based on a CurA GNAT crystal structure, Thr335 and His389 (CurA numbering, Figure 2.4B) in the active site were proposed to mediate only decarboxylation. Substitutitons to these residues did not affect the loading activity.

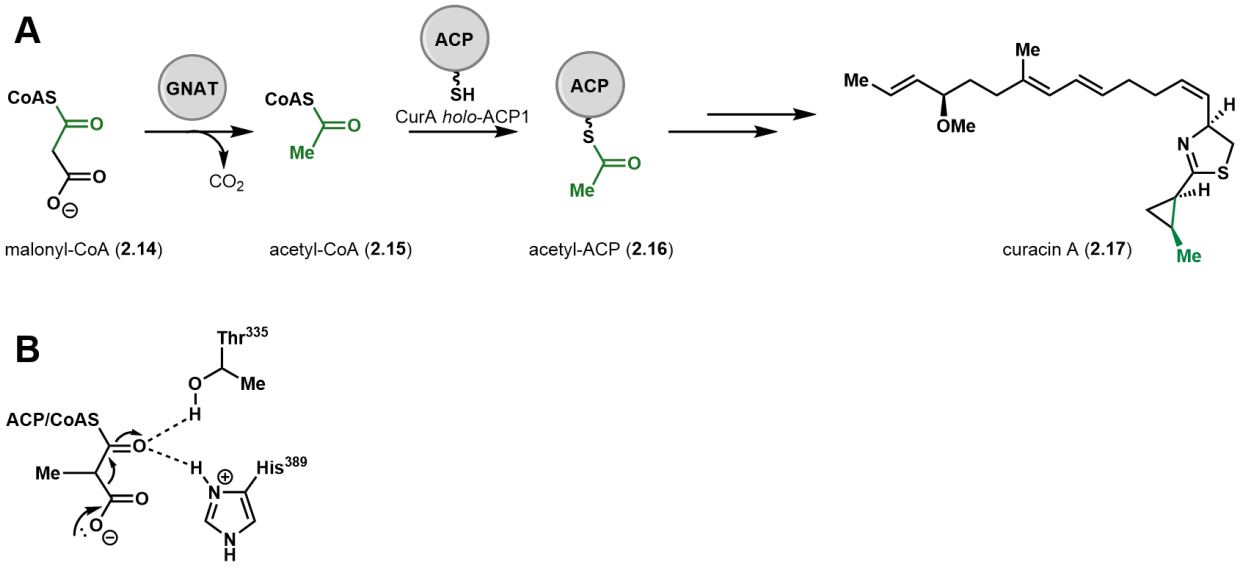


Figure 2.4. GNATs in polyketide synthases.
(A) Initially reported loading and decarboxylation activity of the CurA GNAT domain. (B) Proposed mechanism of CurA GNAT-mediated decarboxylation.[25]

The reports of CurA GNAT activity and the *C. raciborskii sxt* gene cluster were published at approximately the same time. In accordance with the original GNAT name, the Neilan lab proposed that SxtA GNAT catalyzed acetyl group transfer from the CoA thioester to *holo*-ACP, although we also surmised that the domain could also use malonyl-CoA as a starter unit in the same fashion as CurA GNAT (see **2.14** to **2.26**, Figure 2.6).

The first domain of SxtA, methyltransferase (MT), has the highest resemblance to MT_L domains adjacent to GNATs (both complete and truncated)[26,27] within PKS loading modules. These domains are distantly related to general PKS MTs and were not predicted to possess the [4Fe–4S] clusters in radical SAM enzymes.[28] Because SxtA MT reacts at a partially activated position alpha to a carbonyl, we assumed that it and other MTs in this class operate with a two-electron mechanism (Figure 2.5A).

Other pathways containing GNAT-associated MT, also known as MT_L domains,[27] include myxovirescin A,[29] gephyronic acid,[30] bryostatin[31] and apratoxin A[32] (Figure 2.5B). Based on the final natural product structures, some of these MT domains were hypothesized to methylate acetyl-ACP or malonyl-ACP twice.[30,33] At the time of our first reactions with SxtA, the methylation activity of MT_Ls had not been recapitulated *in vitro*. From the acetyl-ACP product of the GNAT domain, Neilan and coworkers proposed that SxtA MT methylates at the α-carbon to afford propionyl-ACP (see **2.26** to **2.27**, Figure 2.6).

As we found more GNAT literature that had been published after the *C. raciborskii sxt* gene cluster report, we found less favor with the hypothesis that acetyl-CoA (**2.15**) is the SxtA starter unit, or that acetyl-ACP (**2.26**) is the MT substrate (Route 1 in Figure 2.6). Instead, malonyl-CoA (**2.14**) appeared increasingly likely to be the true starter unit, and malonyl-ACP (**2.28**) to be the

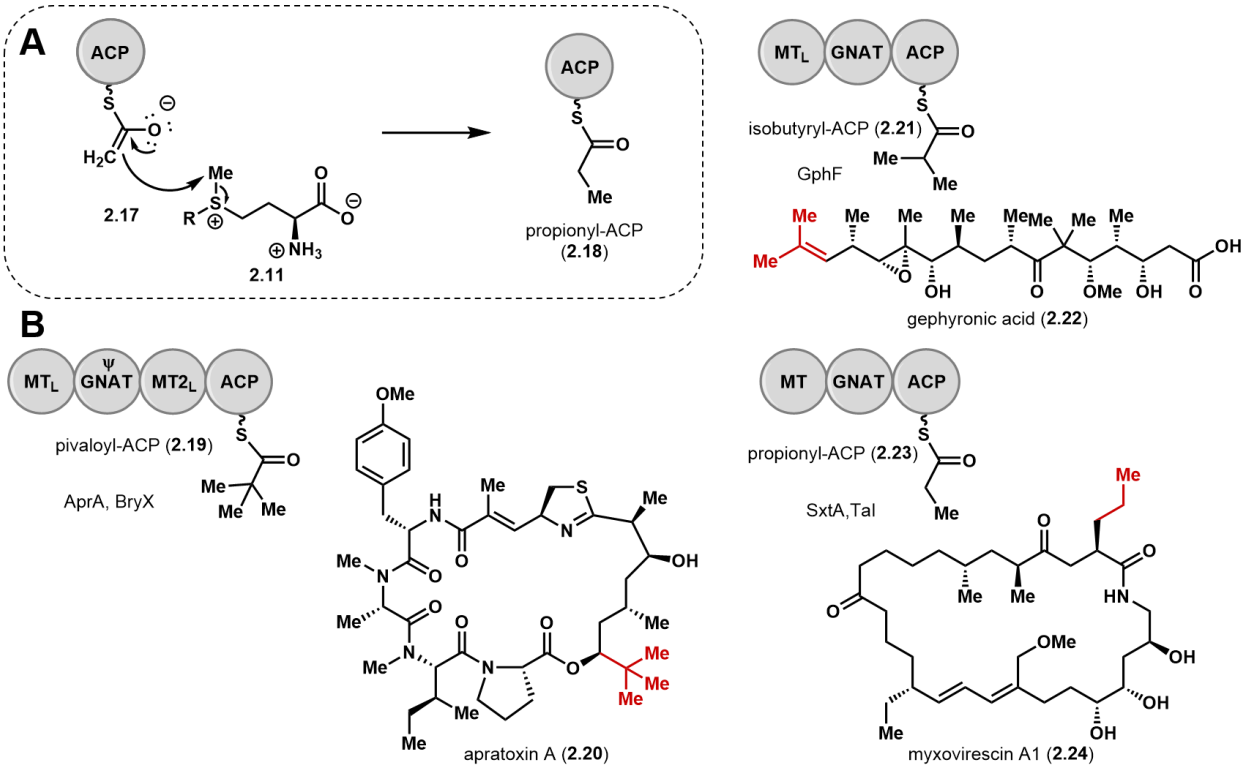


Figure 2.5. GNAT-associated MTs (MT_Ls) in natural product biosynthesis.
(A) Putative methylation mechanism. (B) A selection of mono-, di-, and trimethylated acyl-ACP intermediates with final polyketide products.[26,29,34]

MT substrate, as Taylor and Keatinge-Clay have previously suggested.[30,33] First, SxtA's second domain is not annotated as a typical PKS acyltransferase (AT) domain which can simply load an acetyl group from acetyl-CoA; rather, the domain is related to the PKS GNATs which were reported to favor malonyl-CoA substrates.[25] Second, Route 1 proposes that both decarboxylation and loading activities by GNAT occur before the MT, whereas partner enzyme (PE) domains in type I PKS typically catalyze their reactions in order from the N-terminus to the C-terminus.[35] Finally, because SxtA MT presumably deprotonates the α-carbon of the acyl-ACP substrate to generate a nucleophile that can attack the SAM methyl group (Figure 2.5A), we also considered the pK_a at that site. The α-carbon of acetyl-CoA with a pK_a around 19-20 can be deprotonated under biological conditions (cf. citrate synthase),[36] but the value at the analogous position is several units lower in malonyl-CoA/-ACP.[37] A second route of MT-GNAT activity (Figure 2.6)

could use a malonyl-CoA starter unit and *holo*-ACP by the GNAT domain or another acyltransferase protein. Malonyl-ACP **2.28** is a better substrate for the MT domain and the next intermediate, methylmalonyl-ACP **2.29** could be decarboxylated by the GNAT domain, which catalyzed decarboxylation faster than apparent loading in CurA. Methylmalonyl-ACP **2.29** is also a β-keto acid which could readily decarboxylate spontaneously, meaning it is likely that the AONS thioester substrate is still propionyl-ACP (**2.27**). Except for GNAT-mediated loading, Route 2 circumvents the concerns about the starter unit and relative reactivity we had with Route 1.

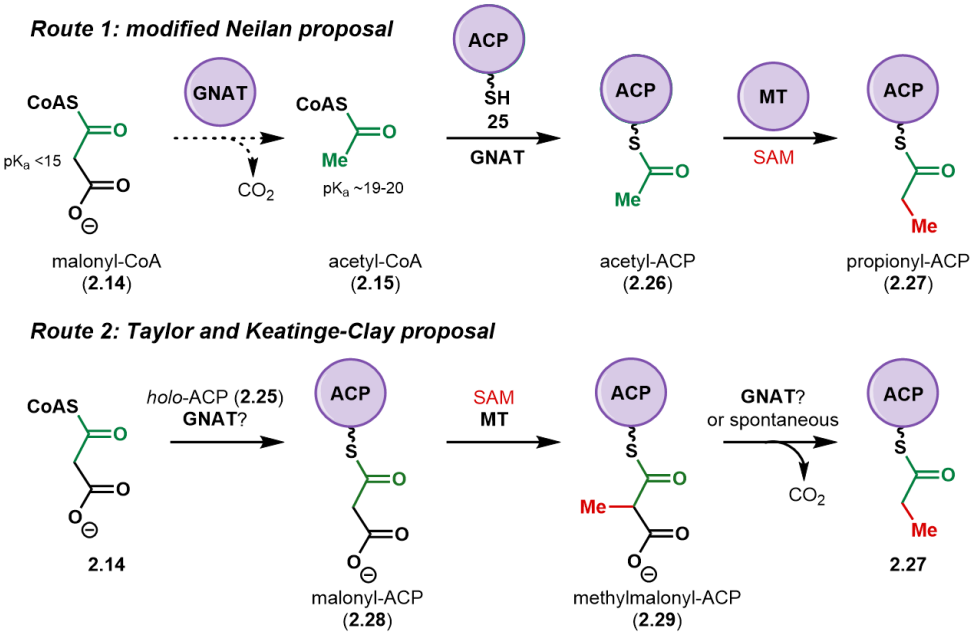


Figure 2.6. Two proposed routes to propionyl-ACP (2.27).

SxtA ACP did not appear to have any unusual features. The protein sequence SxtA AONS' is most closely related to a protein from biotin biosynthesis, 8-amino-7-oxononanoate synthase or BioF, from which the domain takes its name.[18,38] SxtA AONS and BioF belong to the small α-oxoamine synthase (AOS) class of pyridoxal phosphate (PLP)-dependent enzymes. These proteins perform a decarboxylative Claisen-like condensation between an α-amino acid and an acyl thioester to afford an α-amino ketone, carbon dioxide and a thiol.[39] The proposed mechanism

begins with the internal aldimine intermediate (**2.30**, see Figure 2.7) with PLP covalently bound to a catalytic lysine residue (Lys1077 in *C. raciborskii* SxtA). An amino acid substrate such as L-Arg (**2.9**) in SxtA displaces the lysine residue to form external aldimine I (**2.31**), which now has a much more acidic proton alpha to the carboxylate group. Deprotonation by lysine or another base affords the nucleophilic quinonoid intermediate I (**2.32**), which attacks the thioester partner (propionyl-ACP in SxtA). Collapse of the tetrahedral intermediate eliminates *holo*-ACP to give a 1,3-dicarbonyl (**2.35**) that is easily decarboxylated and reprotonated stereospecifically (see **2.36** to **2.37**). Displacement of the ketone product **2.13** returns to the resting state. In 2016, there were five other ACP-associated AOS family members, six that used CoA thioester substrates, and two with either ACP or CoA substrates.

Our synthetic interest in SxtA was primarily in the AONS domain, which forms a new C–C bond, retains the L stereocenter through stereospecific proton transfer steps to synthesize α-amino ketones in a single step, and also functions as SxtA's offloading mechanism. Furthermore, SxtA-mediated synthesis of **2.13** analogs could serve as precursors to STX derivatives.

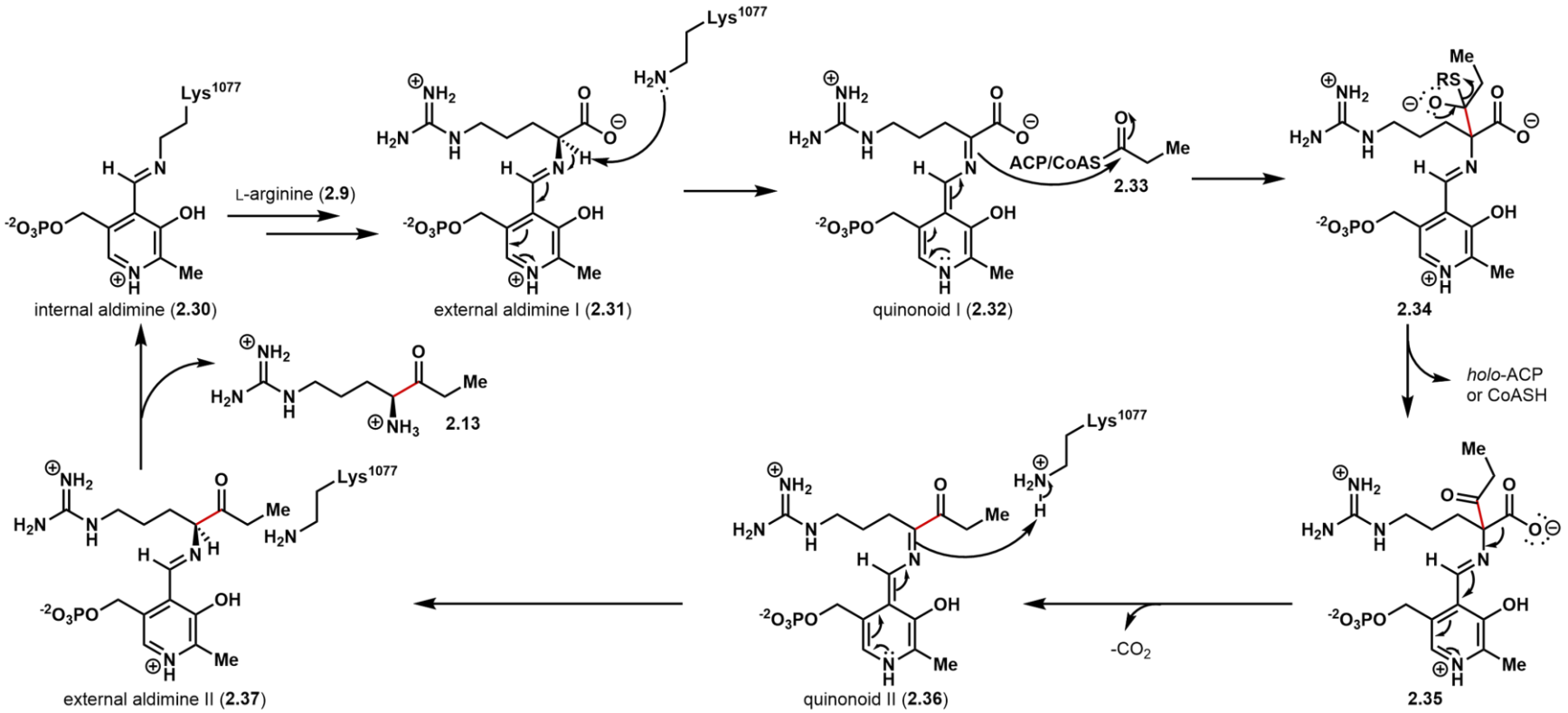
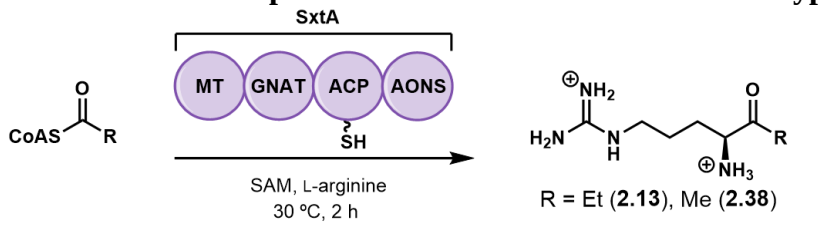


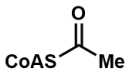
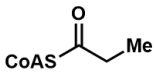
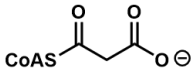
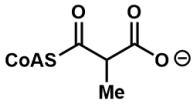
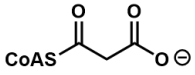
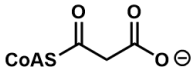
Figure 2.7. Proposed mechanism of PLP-dependent AOS proteins.
Based on Webster et al.[39]

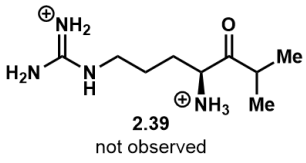
2.3 Enzymatic Reactions with the Wild-Type SxtA Module

sxtA from *C. raciborskii* T3 (gDNA sequence) cloned into a pET28 vector was transformed into the *E. coli* strain BAP1, which the Khosla lab engineered to genomically encode an inducible copy of the phosphopantetheinyltransferase *sfp*.[40] The protein expressed well, but the majority was lost with the insoluble material during purification. After FPLC purification on Ni-NTA resin, a pale-yellow protein of approximately 140 kDa was isolated. Although SxtA did not ionize sufficiently to analyze by intact protein MS, we assumed complete post-translational modification of *apo*-ACP to *holo*.

Table 2.2. Screen of potential CoA starter units with wild-type SxtA



entry	*substrate*	*SAM*	*Arg*	*normalized product yield*	*% ethyl ketone (2.13)*	*% methyl ketone (2.38)*
1		+	+	91 ± 7	ND	100
2		+	+	656 ± 40	100	ND
3		+	+	100 ± 15	45	55
4		+	+	519 ± 64	100	ND
5		–	+	165 ± 6	ND	100
6		+	–	4 ± 1	100	ND



2.39
not observed

We first attempted to determine the CoA starter unit for STX biosynthesis from a candidate pool of acetyl-, propionyl-, malonyl-, and methylmalonyl-CoA thioesters. Purified SxtA was incubated with SAM, L-Arg and each of the four thioesters for 2 h before analysis of the small molecule products by liquid chromatography-mass spectrometry (LC-MS). With an acetyl-CoA starter unit, we observed none of the expected product (Table 2.2, entry 1), but instead another peak was detected corresponding to 173.14 m/z, a difference of 14 m/z or one methylene group from the predicted ketone product at 187.15 m/z. This lower mass matched the structure of methyl ketone **2.38**, which could form by an acetyl group bypassing the MT domain and being processed directly by the AONS domain. The alternate product was confirmed as **2.38** by comparison to an authentic standard synthesized by Dr. Meagan Hinze. If this were an on-pathway intermediate to STX, the terminal methyl group would need to be extended by one carbon by a downstream enzyme such as the putative standalone methyltransferase SxtX.[18] We briefly considered the idea that SxtA MT may be vestigial or currently inactive, as solely the originally predicted ethyl ketone **2.13** was detected when the longer propionyl-CoA thioester was used as a possible CoA starter unit. Like the acetyl substrate, the propionyl group only underwent AONS-catalyzed condensation, skipping the MT domain that would have resulted in isopropyl ketone **2.39**.

Third, malonyl-CoA, the proposed GNAT domain starter unit, was tested. In these reactions, a mixture of the truncated methyl (**2.38**) and full-length ethyl (**2.13**) ketones was observed. In the product mixture, 45% had been methylated by the MT domain and decarboxylated at some point en route to **2.13**, while the remaining 55% was only decarboxylated (either spontaneously or GNAT-mediated) and not methylated before the condensation reaction to afford **2.38**. Control reactions with malonyl-CoA but omitting either SAM or L-Arg were also performed (entries 5 and 6). Without SAM, no methylation occurred and only **2.38** was detected, supporting the proposal

that STX C13 is SAM-derived. We observed only the full-length ketone **2.13** even when L-Arg was not added to reactions, likely forming from amino acid that had copurified with SxtA. Fourth, methylmalonyl-CoA was only decarboxylated (entry 4), eventually forming the proposed native product **2.13** and no other ketones with further methylation such as **2.39**.

Among the four CoA thioesters, only malonyl-CoA or an intermediate resulting from it was methylated by the MT domain, suggesting that malonyl-ACP **2.28** was indeed the MT substrate. The product of this reaction is methylmalonyl-ACP **2.29**, which can be decarboxylated by the GNAT domain to propionyl-ACP (**2.27**). The total relative full-length ketone yields in Table 2.2 also support the malonyl- to methylmalonyl- to propionyl-ACP order of intermediates. In this sequence, acetyl-CoA or -ACP does not fit. We were unable to determine the extent of ACP loading with the full SxtA module by intact protein MS. Acetyl thioesters were never methylated by the MT domain under any conditions in our hands to afford ketone **2.13**, and addition of putative methyltransferase SxtX did not increase the yield of **2.13** relative to the truncated ketone **2.38**. Furthermore, **2.38** or STX derivatives missing C13 have not been identified from natural producers. Therefore, acetyl-CoA was excluded as a possible SxtA starter unit, along with methylmalonyl-CoA, which is not produced in cyanobacteria such as *C. raciborskii*.[41] Since propionyl-CoA/-ACP was only processed by the AONS domain and not the MT domain, it was also excluded as the SxtA starter unit, leaving only malonyl-CoA as the only plausible on-pathway substrate. This thioester is biosynthesized from acetyl-CoA in a reaction catalyzed by acetyl-CoA carboxylase,[42] supporting the observation by Shimizu that acetic acid is the source of C5-6 on STX (**2.1**) and that only one deuterium atom is retained from the deuterium labeled acetic acid (see **2.10**, Figure 2.2B).

From this screen of SxtA with four possible starter units, malonyl-CoA **2.14** was identified as the most likely candidate, and a partial sequence of domain activity was determined in agreement with Route 2 (Figure 2.6): (1) loading of the malonyl group from **2.14** onto *holo*-ACP, (2) methylation of malonyl-ACP to methylmalonyl-ACP, (3) GNAT-catalyzed decarboxylation to propionyl-ACP, and (4) AONS-mediated condensation with L-Arg to form ketone **2.13**. The loading mechanism could not be determined with full module experiments and was explored through reactions with standalone SxtA domains.

2.4 Verification of Domain Activity

From the SxtA module experiments, the three catalytic domains, MT, GNAT and AONS appear to be active, but we could not analyze any ACP-bound intermediates due to the size of the protein. Standalone single and didomains were first subcloned from the wild-type SxtA module by ligation-independent cloning,[43] then expressed and purified on Ni-NTA resin to verify the activity of each domain.

Potential acyl-CoA species present are difficult to detect on the mass spectrometers available to us but could be separated and analyzed by a relatively low-throughput liquid chromatography method (66 min per sample).[26] However, at 13.5 kDa, SxtA ACP constructs ionize very well compared to the full module. ACP-bound intermediates are analyzed more rapidly by intact protein MS (15 min per sample) and the secondary Ppant ejection assay, an MS/MS method that detects the Ppant cofactor, along with any loaded acyl groups.[44]

The Smith lab had previously found that all constructs of AprA MT_L,[26] a homolog of SxtA MT, were insoluble, so we cloned SxtA MT-GNAT instead, which resulted in soluble and well-behaved protein. Other constructs included SxtA GNAT, ACP (*apo* and *holo*), AONS and

ACP(knockout)-AONS. Like the full module, SxtA AONS and ACP(knockout)-AONS had high expression but a significant amount was lost in the insoluble pellet during purification. The didomain that included an inactive ACP was slightly more soluble. Purified SxtA GNAT tended to precipitate at temperatures above rt.

Additionally, single-point variants of the SxtA module were also prepared, each removing or reducing activity in one domain by substituting proposed catalytic residues in the active sites: S773A (ACP knockout), H372A (in the MT domain), T637V, H671A (GNAT residues), and K1077L (AONS substitution).

Loading of Malonyl-CoA

The loading mechanism of malonyl-ACP formation from malonyl-CoA and *holo*-ACP was examined first. From the CurA GNAT precedent, SxtA GNAT was still our initial hypothesis for the loading enzyme. When malonyl-CoA, holo-ACP, and SxtA GNAT or MT-GNAT were incubated together, a mixture of malonyl- and acetyl-ACP was observed (Figure 2.8A) by intact protein MS. At early timepoints, the loading levels appeared to be dependent on the concentration of GNAT constructs present and plateaued around 80% loading after 2 h reactions. To our surprise, loaded ACP was also detected in the negative control in the absence of GNAT or MT-GNAT. Loading reached 80% in 2 h, compared to 1 h or less in the MT-GNAT/GNAT-containing reactions.

After searching the PKS literature further, we found that other laboratories had observed this apparent self-loading for type II ACPs *in vitro* with malonyl-CoA with two potential explanations for the mechanism. Many type II PKS gene clusters do not encode an acyltransferase (AT) protein to transfer malonyl extender units from malonyl-CoA, leading some, including Crosby and

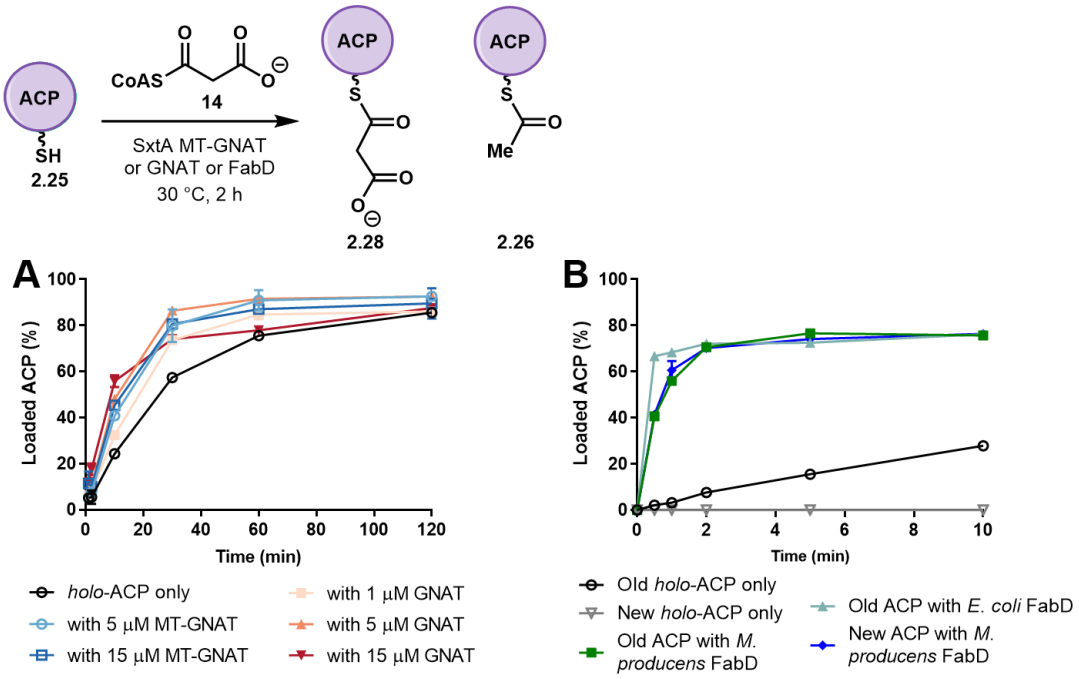


Figure 2.8. Loading of *holo*-ACP from malonyl-CoA.
Formation of malonyl- and acetyl-ACP after incubation of *holo*-ACP, malonyl-CoA with (A) GNAT constructs or (B) FabD.

Simpson labs to postulate that these pathways function by ACP self-loading.[45,46] Other labs such as the Reynolds and Khosla groups eliminated the "self-loading" activity of type II ACPs after further purification by anion exchange chromatography. They proposed that AT-less pathways co-opt FabD, a malonyl-CoA:ACP transacylase protein from fatty acid biosynthesis instead.[47,48] *In vitro* cascade reactions with extra pure ACPs and their associated PEs required the addition of nanomolar concentrations of FabD to proceed; therefore, the apparent self-malonating activity was a side effect of miniscule amounts of endogenous host *E. coli* FabD or some other acyltransferase contaminating the recombinant enzymes of interest.[47] Indeed, SxtA *holo*-ACP that was more rigorously purified by nickel affinity, anion exchange and size-exclusion chromatography no longer exhibited the self-loading activity *in vitro* (Figure 2.8B). In the presence of 50 nM *E. coli* FabD or *Moorea producens* FabD (kindly provided by Dr. Meredith Skiba of the Janet Smith lab),

100 μM *holo*-ACP and 500 μM malonyl-CoA, the ACP now reached 80% loading within 5 minutes. Including FabD with a full SxtA reaction was also favorable, as its presence changed the distribution of products to 93% of full-length ethyl ketone **2.13** and 7% truncated ketone **2.38**.

Finally, after publication of our initial SxtA manuscript,[49] the Smith lab also cultured *E. coli* containing an empty expression plasmid and enriched for divalent cation-binding proteins by Ni-NTA affinity purification. The mixture of endogenous *E. coli* proteins isolated catalyzed acylation of CurA *holo*-ACP but did not effect the decarboxylation of malonyl-ACP to acetyl-ACP.[50] From these data, SxtA is likely natively loaded by FabD or a similar acyltransferase, not its GNAT domain. The apparent GNAT-mediated loading or self-loading of malonyl groups onto SxtA *holo*-ACP is the result of contaminating *E. coli* proteins co-purifying with our proteins of interest, while the GNAT domain's principal function is decarboxylation. The 2013 discovery that malonyl-CoA decarboxylases have the same fold as acyl-transferring GNAT proteins also support this proposal.[51,52]

Methylation of Malonyl-ACP

Following loading of a malonyl group onto *holo*-ACP, the next steps in our proposed SxtA catalytic cycle are methylation by the MT domain and GNAT-catalyzed decarboxylation. Acyl-ACP substrates to verify MT and GNAT activity were prepared enzymatically from *apo*-ACP, phosphopantetheinyltransferase Sfp and acyl-CoA. Loaded ACP was separated from PPTase and excess CoA by size-exclusion chromatography before use.

To assess the didomain cascade, wild-type MT-GNAT, malonyl-ACP and SAM were incubated together for 3 h, and the levels of each possible acyl-ACP intermediate (malonyl, methylmalonyl, propionyl and acetyl) were compared by intact protein MS. At early timepoints

(Figure 2.9), methylmalonyl-ACP could clearly be observed, but then the intermediate decarboxylated to accumulate propionyl-ACP over time. Some acetyl-ACP was present in the initial malonyl-ACP sample due to spontaneous decarboxylation, and the relative fraction remained around 20% throughout the experiment. Even when all malonyl-ACP had been consumed, acetyl-ACP was not converted to propionyl-ACP (see 2 to 3 h). Dimethylated species dimethylmalonyl-ACP or its decarboxylated product isobutyryl-ACP were not detected. In control experiments lacking MT-GNAT, we did not observe methylation of malonyl-ACP.

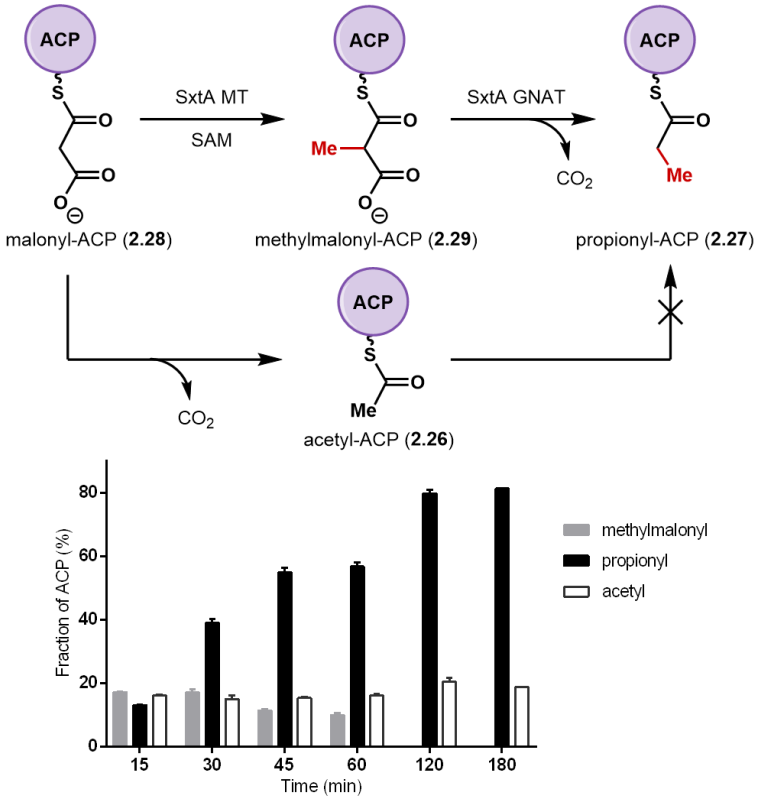


Figure 2.9. Methylation and decarboxylation of malonyl-ACP.

From these data, SxtA MT's function is to methylate malonyl-ACP once to afford methylmalonyl-ACP. The stereochemistry of this reaction can be predicted from the crystal structures of the SxtA MT homolog, AprA MT_L (see Chapter 5),[26] but was not investigated because

α-alkyl-β-keto-ACP intermediates like methylmalonyl-ACP are configurationally unstable and prone to decarboxylation, eroding any stereochemical information.

GNAT with Carboxylated ACPs

The next step to be examined in isolation was the decarboxylation activity of the GNAT domain. The previously generated malonyl-ACP and methylmalonyl-ACP (assumed to be racemic) were incubated with GNAT. Over 3 h, we observed conversion of malonyl- to acetyl-ACP and methylmalonyl- to propionyl-ACP (Figure 2.10). The activity was concentration-dependent, with increasing decarboxylation detected when more GNAT was present. However, it was difficult to determine from this assay alone whether methylmalonyl-ACP was the preferred GNAT substrate. Due to the relative instability of methylmalonyl groups, 54% of acyl-ACP present had already spontaneously decarboxylated propionyl-ACP at 0 min, leaving only 46% of methylmalonyl-ACP substrate for the GNAT to react with, compared to a 19:81 ratio of acetyl-/malonyl-ACP. Additionally, the enzymatic decarboxylation of carboxylated-ACP was also competing with the background spontaneous reactions, which were subtracted from the values shown in Figure 2.10.

Nevertheless, this assay demonstrated that both malonyl- and methylmalonyl-ACP are substrates for GNAT-mediated decarboxylation. We attribute the formation of truncated ketone **2.38** observed previously in the full module reactions to the AONS processing acetyl-ACP/-CoA that arose from a combination of the GNAT reacting prematurely with malonyl-ACP/-CoA and the background nonenzymatic decarboxylation. In the endogenous STX pathway, we hypothesize

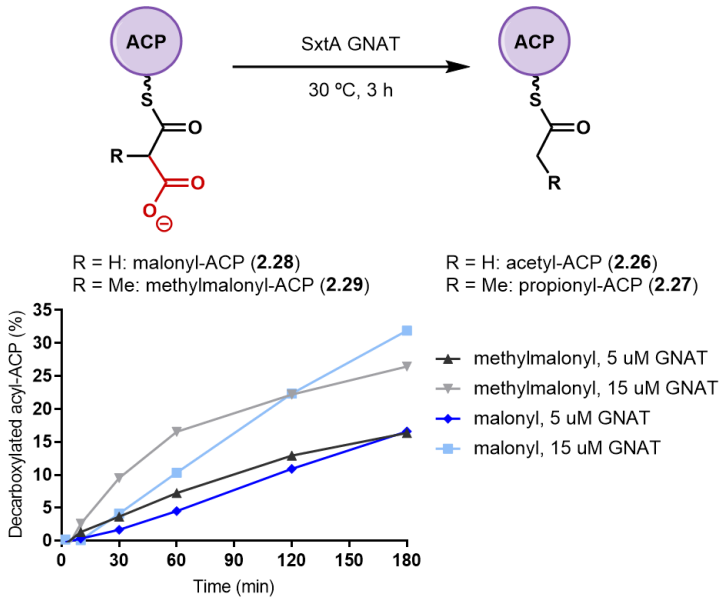


Figure 2.10. GNAT-catalyzed decarboxylation of malonyl- and methylmalonyl-ACP.

that methylmalonyl-ACP is the preferred GNAT substrate over malonyl-ACP, and a competition experiment between malonyl- and methylmalonyl-ACP in the same pot will be tested in the future.

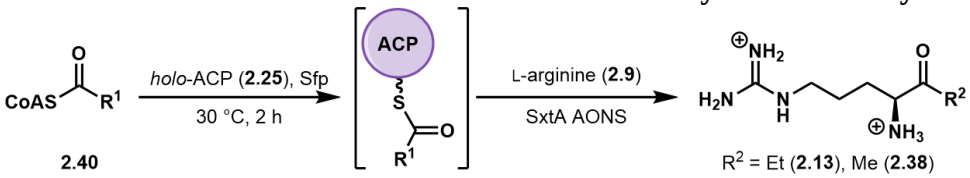
α-Amino Ketone Formation Catalyzed by SxtA AONS

The final step of the proposed SxtA catalytic cycle is the Claisen-like condensation catalyzed by the AONS domain. The *C. raciborskii* standalone AONS was active but not very soluble, so we also cloned a didomain variant ACP(KO)-AONS containing an inactive ACP domain. This construct was subcloned from SxtA(S773A) and was slightly more soluble than the single domain.

We first aimed to compare AONS condensation activity on acyl-ACP intermediates, from the four CoA thioesters in our possession: acetyl-, propionyl-, malonyl-, and methylmalonyl-ACP. L-Arg, acyl-CoA, *apo*-ACP, Sfp and ACP(KO)-AONS were incubated in a single pot (Table 2.3). Small molecule analysis after 2 h showed the same two methyl and ethyl ketones (**2.38** and **2.13**, respectively) we had observed previously, and no carboxyl ketones. The relative yield was highest with a propionyl-CoA substrate, followed by methylmalonyl, acetyl, and just a small amount of

2.38 from malonyl-CoA. Yields of the alkyl substituents were higher than the corresponding carboxylated substrates, suggesting that SxtA AONS does not process or strongly disfavors the bulkier charged malonyl- and methylmalonyl-thioesters. If the carboxylated groups were to form α-amino ketones, they would likely decarboxylate spontaneously before we could observe them.

Table 2.3. AONS-mediated ketone formation via acyl-ACP and acyl-CoA thioesters



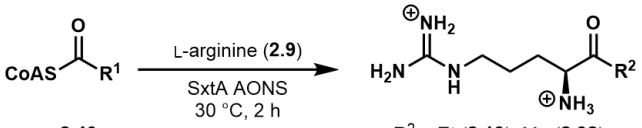
2.40

R^2 = Et (**2.13**), Me (**2.38**)

Entry	Thioester Substrate	Product	Relative Yield
1	acetyl-CoA (2.15)	2.38	22 ± 3
2	propionyl-CoA	2.13	100 ± 12
3	malonyl-CoA (2.14)	2.38	1.8 ± 0.1
4	methylmalonyl-CoA	2.13	74 ± 1

Intact protein MS indicated, *apo*-ACP was completely modified to acyl- or *holo*-ACP, but because the CoA thioesters were present in 10-fold excess relative to the concentration of *apo*-ACP, it was possible that the AONS used both CoA and ACP-bound substrates. Therefore, standalone SxtA ACP was eliminated entirely, and the AONS domain tested directly with the same CoA substrates (Table 2.4). In these reactions, propionyl-CoA was the vastly preferred substrate, and low amounts of the other three groups were processed.

Table 2.4. Direct AONS-mediated ketone formation from acyl-CoA thioesters



2.40

R^2 = Et (**2.13**), Me (**2.38**)

Entry	Thioester Substrate	Product	Relative Yield
1	acetyl-CoA (2.15)	2.38	11 ± 1
2	propionyl-CoA	2.13	100 ± 17
3	malonyl-CoA (2.14)	2.38	0.8 ± 0.2
4	methylmalonyl-CoA	2.13	4.6 ± 0.8

SxtA AONS catalyzes the formation of at least two ketones, the predicted native product **2.13**, and a truncated version **2.38**. Due to the higher relative yields of **2.13** over **2.38** in all assays, we propose that propionyl-ACP is the AONS thioester substrate, and that **2.13** is the true native SxtA product.

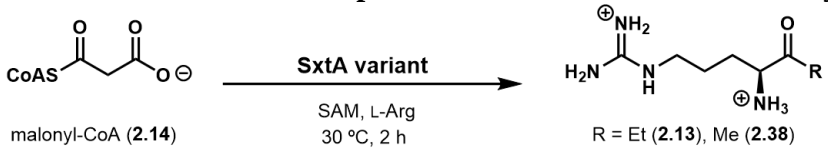
SxtA Single-Point Variants

For final verification of each domain's activity relative to the SxtA module, we assessed five variants of the SxtA module, each with a single substitution in the respective active site. The ACP variant, S773A, removes the serine residue where the Ppant cofactor is attached. This construct was cloned by site-directed mutagenesis of wild-type SxtA and also served as the parent for ACP(KO)-AONS (see above). The other four substitutions in the PE domain active sites were cloned by splicing by overlap extension (SOE) due to the 3.8 kb size of the *sxtA* gene. The His372 residue was implicated in the methyl transfer reaction based on CurJ MT and was changed to Ala.[53] Thr637 and His671 align with the catalytic residues in CurA GNAT and were also substituted. Finally, Lys1077 is proposed to form an imine with the PLP cofactor and may also mediate proton transfers (see Figure 2.7). Although PLP intermediates reside in the AONS active site noncovalently, the resulting SxtA(K1077L) protein was very pale compared to the wild-type enzyme, indicating low or no PLP incorporation. Extra PLP was not added to *in vitro* reactions.

Each variant was incubated with malonyl-CoA, SAM and L-Arg, and the small molecule ketones were analyzed after 2 h. The reactions of wild-type SxtA and the S773A variant (Table 2.5, entries 1 and 2) were similar in total product yield, and in the **2.13**: **2.38** distribution. These reactions indicated that all three PE domains can function on CoA thioesters in addition to ACP-bound thioesters. When standalone *holo*-ACP (entry 3) was added to reactions with SxtA(S773A),

the yield increased, and the ratio of ketone products shifted to favor the full-length ketone **2.13** slightly in a 63:27 ratio of **2.13**/**2.38**. The *trans*-acting ACP used here was not stringently purified with the additional ion exchange step and contains the contaminating acyltransferase(s) that catalyzes formation of malonyl-ACP. Still, the results suggest that the overall SxtA reaction rate is faster on acyl-ACP substrates and that SxtA MT methylates malonyl-ACP more readily than malonyl-CoA. Later, we performed a similar reaction with WT SxtA and *E. coli* FabD as catalysts (entry 8). The complete SxtA reaction dominated, affording a 93:7 ratio of **2.13**/**2.38**.

Table 2.5. Full SxtA reactions with variant proteins to confirm domain activity



malonyl-CoA (**2.14**) R = Et (**2.13**), Me (**2.38**)

entry	*SxtA variant*	*normalized product yield*	*% ethyl ketone (2.13)*	*% methyl ketone (2.38)*
1	MT GNAT ACP AONS SH **WT**	100 ± 18	45	55
2	MT GNAT ACP AONS **S773A**	106 ± 22	42	58
3	MT GNAT ACP AONS + ACP SH **S773A + *holo*-ACP**	387 ± 75	63	27
4	MT GNAT ACP AONS SH **H372A**	46 ± 9	ND	100
5	MT GNAT ACP AONS SH **T637V**	27 ± 1	100	ND
6	MT GNAT ACP AONS SH **H671A**	9 ± 2	100	ND
7	MT GNAT ACP AONS SH **K1077L**	0	ND	ND
8[a]	**WT SxtA with *E. coli* FabD**	N/A	93	7

See section 2.6.IIIA for full reaction conditions. [a]This reaction was performed much later under different reaction conditions than entries 1-7, so the yields were not compared.

Among the PE domain variants, the MT domain of SxtA(H372A) was inactive, resulting in only formation of **2.38** and decreased the overall yield. These values support the proposal that methylmalonyl-ACP/-CoA and propionyl-ACP/-CoA are the preferred substrates of GNAT and AONS, respectively, over the shorter malonyl and acetyl thioesters. Both GNAT variants (entries 5 and 6), SxtA(T637V) and SxtA(H671A), lowered the ketone yield significantly to 27 and 9 relative to the WT reaction, but afforded only full-length ketone. Since previous reactions with the AONS domain indicated that it strongly disfavors carboxylated intermediates, any methylmalonyl-ACP/-CoA must be decarboxylated spontaneously to propionyl thioesters before AONS-mediated condensation. No ketone products were observed when SxtA(K1077L) was the catalyst, and we were unable to determine if the module had any stalled acyl-ACP intermediates by MS.

Reactions with variants of the SxtA module demonstrated that each domain was catalyzing the reactions predicted from the standalone domain activity and provided further evidence of their native substrates.

2.5 Conclusions

From all of the above reactions, we propose the SxtA catalytic cycle outlined in Figure 2.11 to initiate STX biosynthesis, primarily following Route 2 in Figure 2.6. Malonyl-CoA (**2.14**) is the true starter unit of SxtA and is the substrate for an acyltransferase such as FabD to load *holo*-ACP. SxtA MT methylates malonyl-ACP (**2.27**) to methylmalonyl-ACP (**2.28**), followed decarboxylation by the GNAT domain to propionyl-ACP (**27**). AONS-mediated condensation between **2.27** and L-Arg affords the previously predicted native product ketone **2.13**, another molecule of CO_2, and *holo*-ACP, closing the catalytic cycle.

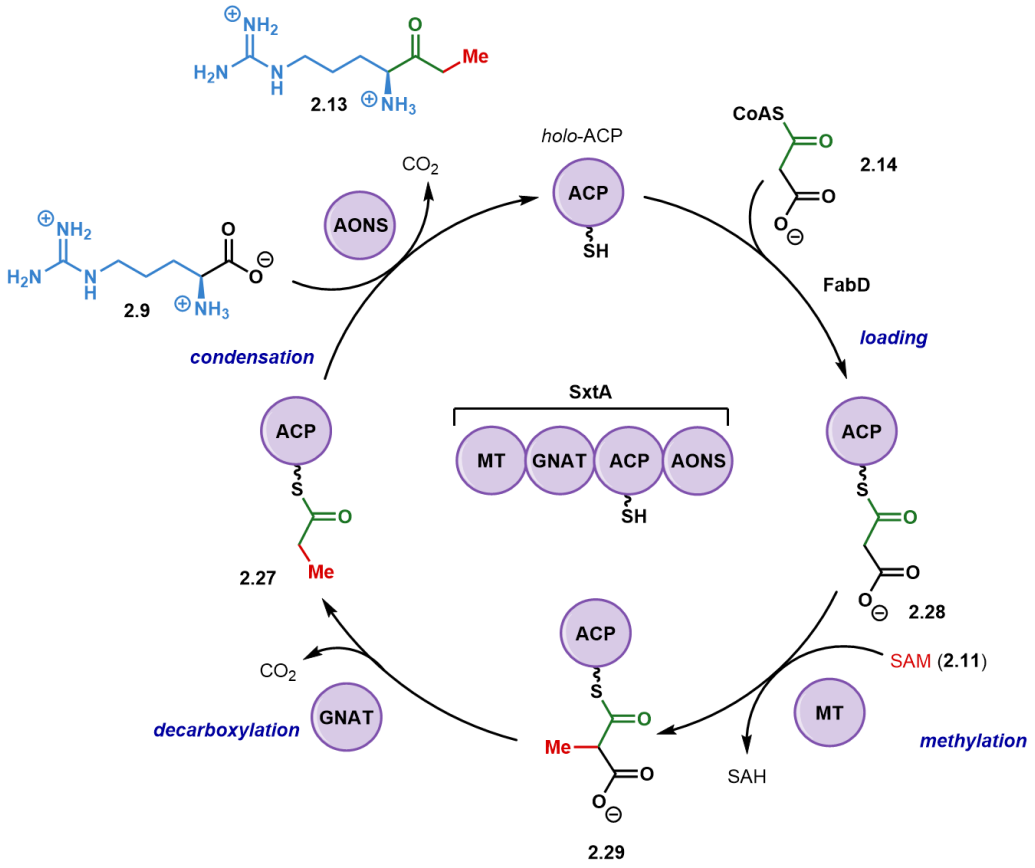


Figure 2.11. The full proposed catalytic cycle of the SxtA module.

Acetyl-CoA is not the SxtA starter unit and acetyl-ACP is not an on-pathway intermediate. Additionally, the truncated ketone **2.38** and STX derivatives lacking C13 that would result from **2.38** have not been reported in the extracts of toxin-producing microorganisms. When the module reactions were slowed in some manner by the exclusion of added L-Arg or impairment of GNAT decarboxylation, only the full-length ethyl ketone **2.13** was detected. Therefore, we hypothesize that methyl ketone **2.38** is an *in vitro* shunt product only.

In this Chapter, we characterized the native activity of the polyketide-like synthase SxtA through bioinformatic sequence analysis of each domain and *in vitro* reactions with the full module and standalone domains. The starter thioester is malonyl-CoA, and the PE domain activity order after loading by some *trans*-acting acyltransferase is MT, then GNAT and AONS to produce

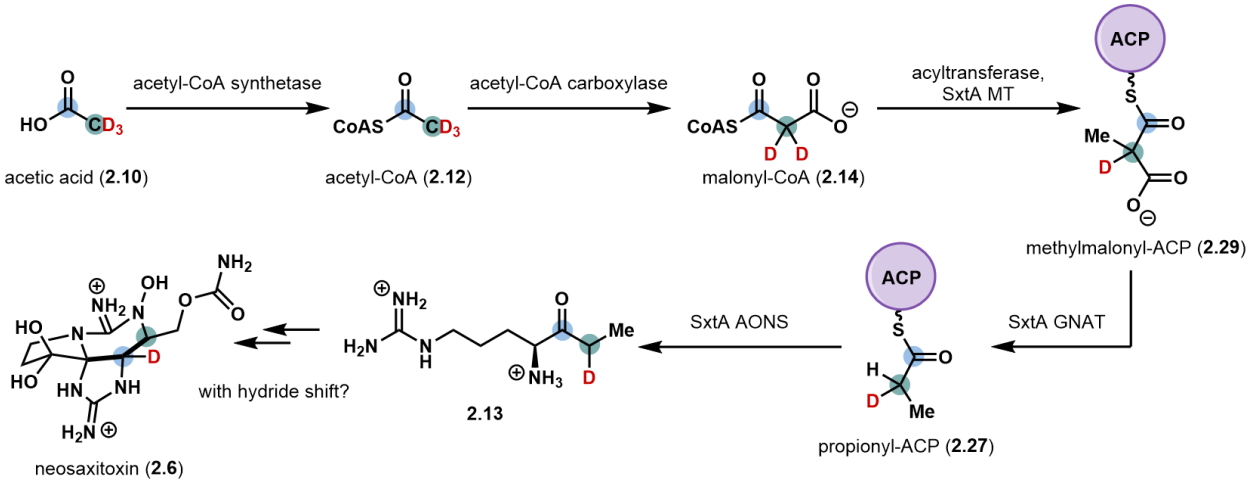


Figure 2.12. Proposed biosynthesis of neosaxitoxin to account for incorporation of one deuterium atom from labeled acetic acid.

ketone **2.13**. These results address our concerns about the pK_a of the MT substrate and the order of domain activity from N- to C-termini and are still in agreement with labeled precursor feeding experiments (Figure 2.12). Our experiments provide the basis for the development of SxtA as a biocatalyst for enzymatic or chemoenzymatical preparation of STX derivatives, and SxtA AONS in particular as a tool for modifying α-amino acids.

2.6 Experimental

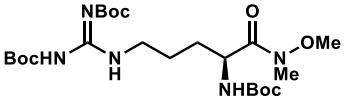
I. Chemical synthesis

A. General information

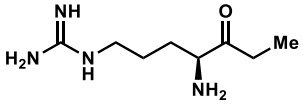
All reagents were used as received unless otherwise noted. Reactions were carried out under a nitrogen atmosphere using standard Schlenck techniques unless otherwise noted. Solvents were degassed and dried over aluminum columns on an MBraun solvent system (Inert Corporation, Model PS-00-3). Trifluoroacetic acid was distilled prior to use. Reactions were monitored by thin layer chromatography using Machery-Nagel 60 F_{254} precoated silica TLC plates (0.25 mm) or Merck Silica Gel 60 RP-18 WF-254S precoated silica TLC plates (0.25 mm) which were visualized using UV, ninhydrin, *p*-anisaldehyde, CAM, DNP, or bromocresol green stain. Flash column chromatography was performed using Machery-Nagel 60 μm (230-400 mesh) silica gel. All compounds purified by column chromatography were sufficiently pure for use in further experiments unless otherwise indicated. ^{1}H and ^{13}C NMR spectra were obtained in $CDCl_3$ or CD_3OD at rt (25 °C), unless otherwise noted, on Varian 400 MHz, Varian 500 MHz or Varian 600 MHz spectrometers. Chemical shifts of ^{1}H NMR spectra were recorded in parts per million (ppm) on the δ scale. High resolution electrospray mass spectra were obtained on an Agilent G6545A quadrupole-time of flight mass spectrometer in positive mode with an Agilent 1290 UPLC system. Solvent A = water with 0.1% formic acid. Solvent B = 95% acetonitrile, 5% water and 0.1% formic acid. IR spectra were recorded on a Perkin-Elmer Spectrum BX FT-IR spectrometer. Optical rotations were measured at rt in CH_3OH, unless otherwise noted, on a Jasco P-2000 polarimeter.

B. Compound synthesis

All compounds in this Chapter were synthesized and characterized by Dr. Meagan Hinze.

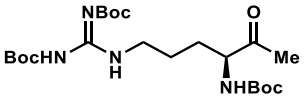


2.S1



2.13

Tri-Boc-protected arginine Weinreb amide (**2.S1**) and arginine ethyl ketone (**2.13**) were synthesized according to the procedure described by Tsuchiya et al.[23]

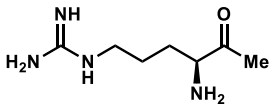


2.S2

Tri-Boc-protected arginine methyl ketone (2.S2):

To a solution of tri-Boc-protected arginine Weinreb amide (**2.S1**, 72.0 mg, 0.139 mmol, 1.0 equiv) dissolved in THF (6.9 mL) was added a 1 M solution of zinc chloride in diethyl ether (13.9 μL, 0.013 mmol, 0.1 equiv) at 0 °C. After five min, a 3 M solution of methyl magnesium bromide in diethyl ether (695 μL, 2.09 mmol, 15.0 equiv) was added at 0 °C. The reaction was allowed to warm to room temperature and stirred for 14 h. The reaction was poured into 21 mL of 0.1 M HCl. The acidified aqueous phase was extracted with EtOAc (3x). The organic layers were combined, washed with brine, dried over Na_2SO_4, filtered, and evaporated under reduced pressure to afford an oil. The crude residue was purified by flash column chromatography (2:1 hexanes/EtOAc v/v) to afford 28.5 mg of the title compound (43% yield) as a white foam. R_f = 0.52 (2:1 hexanes/EtOAc v/v); **^{1}H NMR** (300 MHz, $CDCl_3$) δ 11.46 (s, 1H), 8.33 (t, *J* = 5.4, 1H), 5.36 (d, *J* = 7.6, 1H), 4.35-4.28 (m, 1H), 3.42 (q, *J* = 6.5, 2H), 2.20 (s, 3H), 1.95-1.85 (m, 1H), 1.66-1.52 (m, 3H), 1.48 (s, 9H), 1.47 (2, 9H), 1.42 (s, 9H); **^{13}C NMR** (150 MHz, $CDCl_3$) δ 207.1, 163.6, 156.4, 155.6, 153.4, 83.3, 79.9, 79.4, 59.7, 40.3, 28.5, 28.4, 28.2, 27.1, 25.2; **IR** (thin film, cm^{-1}) 3329, 2978, 1715,

1614, 1498, 1414; **HRMS** (ESI) *m/z* calculated for $C_{22}H_{41}N_4O_7^+$ $[M+H]^+$ 473.2970, found 473.2977.

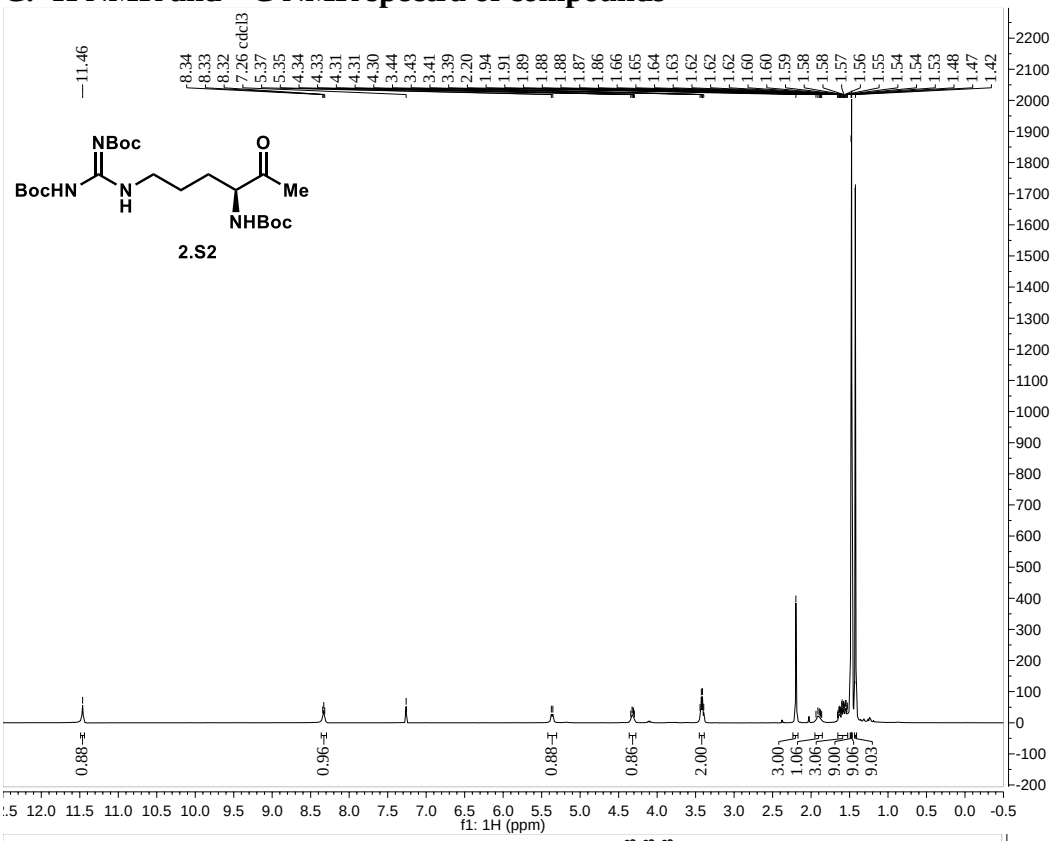


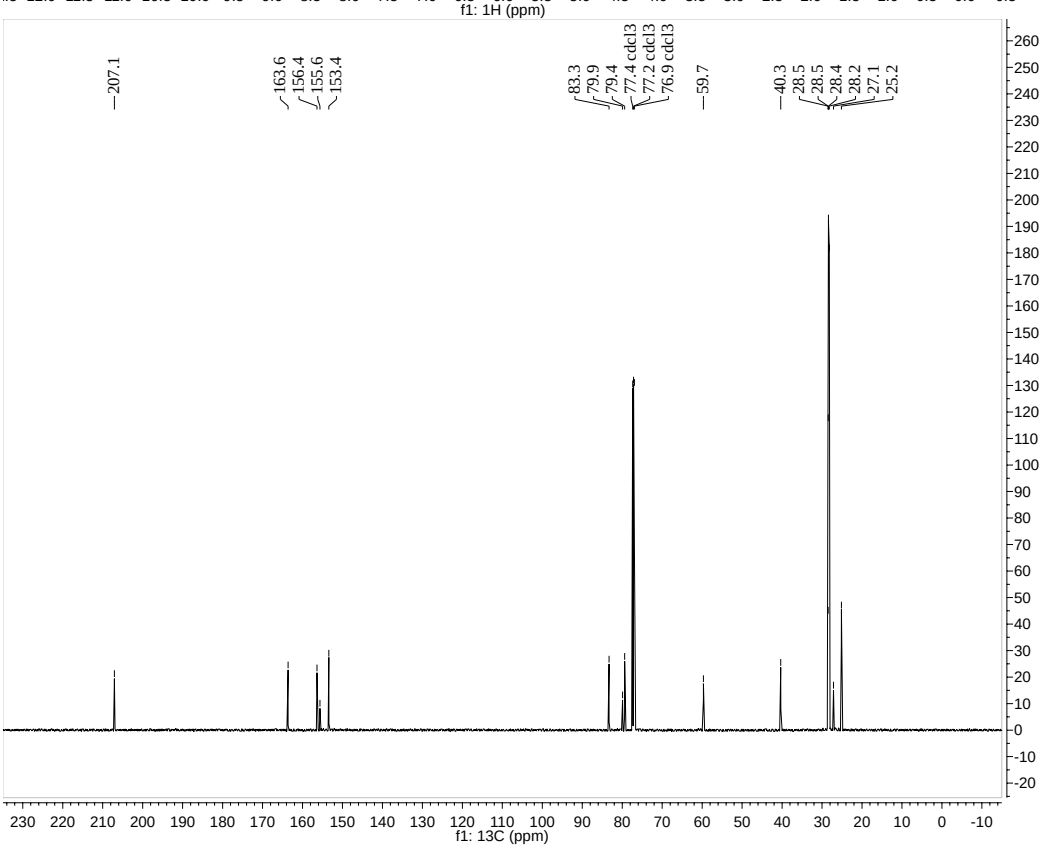
2.38

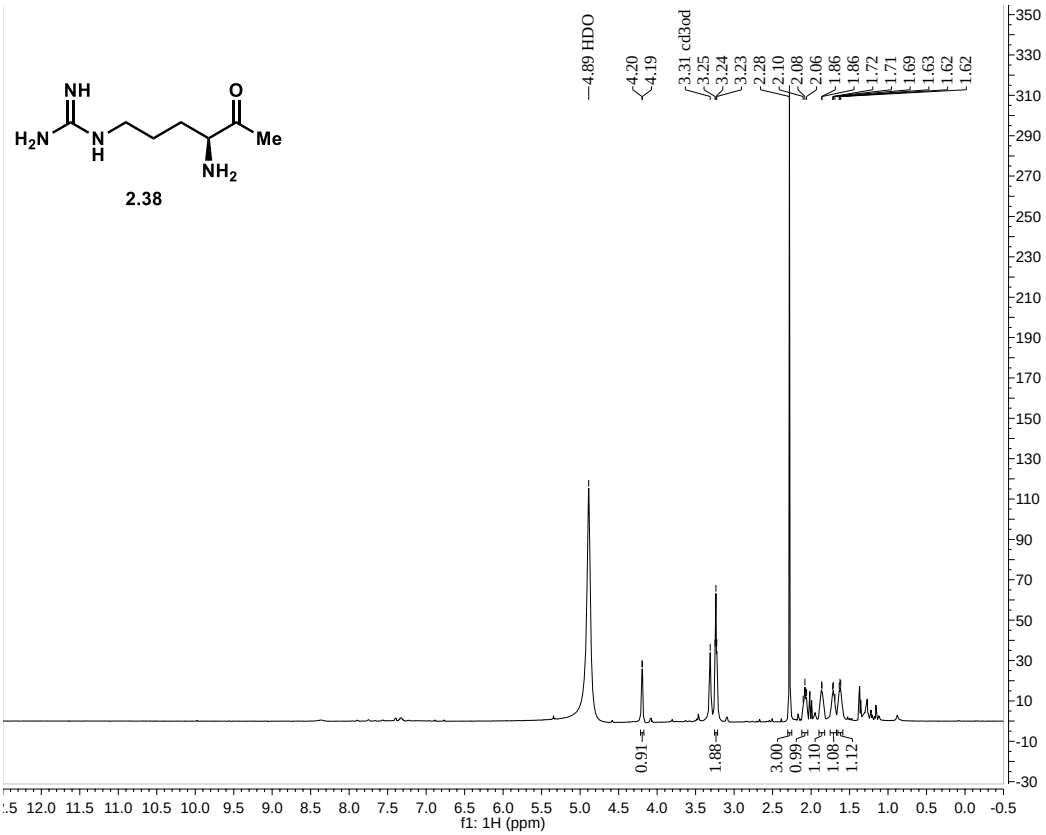
(*S*)-1-(4-amino-5-oxohexyl)guanidine (2.38):

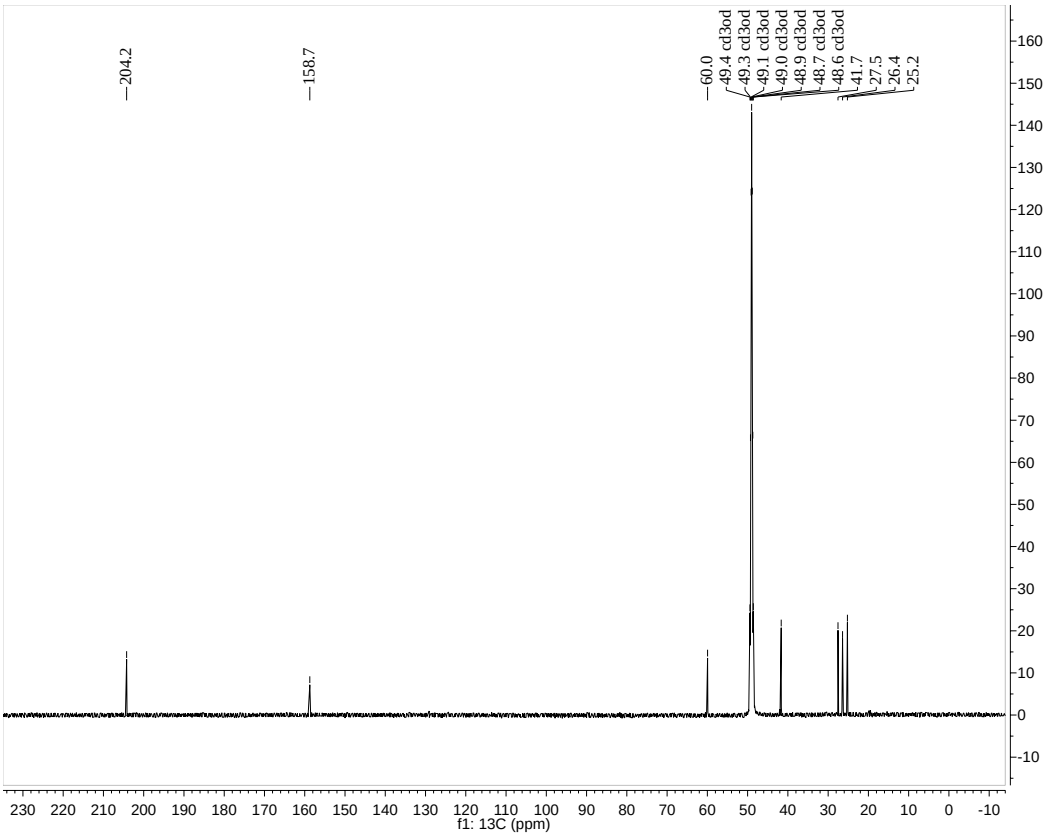
To a solution of tri-Boc protected arginine methyl ketone (**2.S2**, 28.5 mg, 0.060 mmol, 1.0 equiv) dissolved in DCM (1.4 mL) was added trifluoroacetic acid (TFA, 1.4 mL) at rt. The reaction was stirred for 3 h. After addition of toluene (2.0 mL), the solution was evaporated under reduced pressure to afford 24.0 mg of the title compound as a bis-TFA salt with minor impurities (99% yield) as a brown foam. R_f = 0.10 (RP_{18} TLC, 2:1 MeCN/H_2O v/v); $[\alpha]_D^{24}$ = 0.310 (*c* = 0.014 g/mL, MeOH); **^{1}H NMR** (600 MHz, CD_3OD) δ 4.22-4.17 (m, 1H), 3.24 t, *J* = 6.3, 2H), 2.28 (s, 3H), 2.08 (t, *J* = 13.0, 1H), 1.91-1.81 (m, 1H), 1.76-1.67 (m, 1H), 1.67-1.57 (m, 1H); **^{13}C NMR** (150 MHz, CD_3OD) δ 204.2, 158.7, 60.0, 41.7, 27.5, 26.4, 25.2; **IR** (thin film, cm^{-1}) 3363, 3187, 1670, 1199, 1133; **HRMS** (ESI) *m/z* calculated for $C_7H_{17}N_4O^+$ $[M+H]^+$ 173.1397, found 173.1398.

C. ^{1}H NMR and ^{13}C NMR spectra of compounds






NH
O
H2N
N
H
Me
NH2
2.38
4.89 HDO
4.20
4.19
3.31 cd3od
3.25
3.24
3.23
2.28
2.10
2.08
2.06
1.86
1.86
1.72
1.71
1.69
1.63
1.62
1.62
0.91
1.88
3.00
0.99
1.10
1.08
1.12
f1: 1H (ppm)


204.2
158.7
60.0
49.4 cd3od
49.3 cd3od
49.1 cd3od
49.0 cd3od
48.9 cd3od
48.7 cd3od
48.6 cd3od
41.7
27.5
26.4
25.2
f1: 13C (ppm)

II. Cloning, protein expression and purification

A. General information

Escherichia coli cloning strains DH10B or DH5α (Invitrogen) were used for DNA propagation. Proteins were expressed in *E. coli* strains BL21(DE3) (New England BioLabs), BL21(DE3) pRARE (customized for streptomycin and spectinomycin resistance by the University of Michigan Center for Structural Biology) or BAP1[40] after transformation with genes subcloned into pET24b, pET28b (Novagen) or pMCSG7 (provided by the University of Michigan Center for Structural Biology) plasmids. Cells were grown in either Luria-Bertani broth (LB) or Terrific broth (TB) supplemented with the corresponding antibiotics kanamycin (50 μg/mL), ampicillin (100 μg/mL) or spectinomycin (80 μg/mL) purchased from Gold Biotechnology.

All primers were purchased from Integrated DNA Technologies (IDT). Phusion HF polymerase and DpnI restriction enzyme were purchased from New England BioLabs. Ligation-independent cloning (LIC)-qualified T4 DNA polymerase was purchased from EMD Millipore. QIAquick PCR purification, gel extraction and miniprep kits were purchased from Qiagen. HisPur nickel-nitrilotriacetic acid (Ni-NTA) resin was purchased from Thermo Scientific. Fast protein liquid chromatography was conducted on a GE Healthcare ÄKTA Pure FPLC with 5 mL HisTrap or HiPrep 16/60 Sephacryl S-200 High Resolution columns (GE Healthcare). Proteins were concentrated using Amicon centrifugal filters purchased from EMD Millipore at 4,000 x *g*, 4 °C. PD-10 desalting columns were purchased from GE Healthcare. Protein samples were analyzed on Mini-PROTEAN TGX Gels (4-15%) from BioRad and visualized with Protein Ark Quick Coomassie Stain from Anatrace. Proteins were quantified with the Pierce 660 nm Assay Reagent from Thermo Scientific.

B. Sequence information

A pET24b vector encoding the gene for the promiscuous phosphopantetheinyl-transferase Sfp[54] (*Bacillus subtilis*) was kindly donated by the group of Prof. David Sherman (University of Michigan Life Sciences Institute). GenBank accession: BAA09125.1.

pET28b-*sxtA* (*Cylindrospermopsis raciborskii* T3) was obtained from the laboratory of Prof. Brett Neilan (University of Newcastle).

A gene fragment of *E. coli* FabD (GenBank accession: AAA23742.1) was purchased from IDT.

pET28b-*sxtA* genomic DNA sequence

(GenBank accession ABI75094.1; linkers and 6x-His tags are shown in red)

atgggcagcagccatcatcatcatcacagcagcggcctggtgccgcgcggcagccatATGTTACAAAAGATTAA
TCGTTATACTCACGGCTTTGTGGCGGTTCCCGTTATTCTTGCGTGTCGAGAAAAAGGTGTTTTTGAATTACTCGCCG
ATGAAAGTCCTCTCTCTTTAAACCAAATGGTGGAACATCTGGGAGCTAACAGCGGACATTTTCAAGTTGCTTTGAGG
ATGCTCGAGTCTTTACATTGGCTTTCCCGAAATAAGGAGCTTAAATACTCTCTAACCGCAGAAGCAGCGATTCACAA
CAAAATTTCGGAAGACATTCTTCAATTGTACAACCTACCAATTCAGTCTTATTTAGAAGGGAAACAAGGAAATTTGC
TGGGAAGATGGATTGAGCGTTCTTGCCAATTGTGGAACCTGGACAATCCCCTAATGGCAGATTTTTTAGATGGATTA
CTGGTCATCCCACTCCTGCTGGCACTGCACAAACACAACTTGCTTGCAGATTCGGAGGACAAACCTTTGCTCTCCTC
ATTAAGCTCAACAGTGCAAGAAGAGTTGGGTAAGTTATTTCTCCACCTTGGCTGGGCTGACCTTACAGCAGGTCGTT
TGACCATAACCGAACTTGGTCGATTTATGGGAGAGCGAGCCTTGAATACAGCCATAGTGGCGTCCTACACTCCTATG
TTGTCCCGCATTCATGATGTATTGTTTGGCAATTGTCTCTCCGTATTCCAAAGAGATGCTTCCGGTCACGAAAGGCA
CATTGATCGCACCCTTAACGTGATCGGGAGTGGATTTCAACACCAGAAATATTTTGCCGATTTAGAAGAAAGTATCC
TCTCAGTATTCAATCAGTTGCCATTAGAAGAACAACCCAAATACATTACTGACATGGGGTGTGGCGATGGAACTCTC
CTGAAACGAGTCTGGGAAACCATTCAATTTAAGTCTGCTAGGGGAAAAGCACTCGAACAGTATCCCCTGCGTCTTAT
AGGTGTAGATTATAACGAAGCTTCTTTAAAAGCTACCACACGCACCCTTGCTAGCCTTCCCCACTTAGTTTTACAGG
GAGATATTGGGAACCCAGAACAAATGGTGCGTTCTTTAGAAGCTCATGGCATTCATGATCCCGAAAATATCCTGCAC
ATCCGTTCGTTCCTCGACCATGATCGTCTCTTTATTCCTCCTCAGAAAAGAAACGAATTGAAAGAACGTGCTCACTT
ACCTTACCAATCAGTCTGTGTCGATGATCAAGGAGAGCTTATTCCTCCTCATGTTATGGTGCAAAGTTTGGTGGAAC
ACTTAGAAAGATGGTCTCAAGTGGTCAATAAACACGGTTTAATGATTTTGGAGGTCCATTGTTTGGAACCAAGGGTA
GTCTATCAGTTTTTAGACAAAAGCGAAAACTTACATTTCGATGCGTTTCAGGGATTTTCTCAGCAGTATCTTGTGGA
AGCTGAGGTTTTTCTCATGTCTGCTGCACAAGTAGGTCTATTTCCAAAACTAGAGCTTTCTAAAAGATACCCAAAAA
CATTTCCTTTTACTCGCATTACGCTTAATTACTTCGAGAAAAGACCTTACAAAATTAGTCATGCCTATTTGTCAGAT
TTACCTGCCTTAGTTGACTTGGAGGTCAAGTGTTGGCCAGAAAATTTACGGGCATCTACTCATGAAATTCGGCGAAG
ACTTGAGCTAAACCCGCAAGGTAATTTAGTGCTGATTATAGAAGATCAAATTATTGGTGCGATTTATTCCCAAACAA
TTACCAGCACTGAGGCATTAGAGAATGTAAAATATGCGCAAGTGCCGACGTTACATACTCCCCAAGGGTCAGTTATT
CAACTGCTCGCACTAAATATTCTACCTGAGTTTCAGGCGCGGGGGTTAGGAAATGAATTGCGGGACTTTATGCTTTA
CTACTGTACCCTGAAAGGCGGCATTGAGAGCGTGGTGGGTGTAACTCGCTGTCGAAATTATGTCAATTATTCCCAAA
TGCCGATGATGGAGTATTTAAAGCTACACAATGAGCAACGACAGCTTCTGGATCCAATTGTGGGTTTCCATGTGTCG
GGAGGAGCCGAAATTAGGGGAATTATTGCTAATTATCGTCCGGAAGATACAGATAATCTCGGCATGGGTATTTTGAT
TGAATATAACCTGCGCGACAGTGCTTTGCACTCGCCTGGTGATCGCAAAGGACCGTATATTAACTCAGCAATTGGTT
CATTGGTACCAAAAGCAACATCTGCAACTAAGGAAAACAAAACTGTAGCGGATCTCGTTAAAGAATGCATCTTAAAA
GTAATGGGTTCCCAACGTCAGGCAGCCTACGCTCCACAACAAAAACTGCTGGATATGGGATTAGATTCTTTAGATTT
ATTAGAACTGCAAACGCTCCTAGAGGAACGTTTAGGGATCAATCTGTCTGGAACGTTCTTTTTACAAAAGAACACTC
CAACTGCCATCATCACTTATTTCCAAAACCAAGTGGTACAAGAGAAACAATCTGATCTAGCTCCACCTGTTGACTCA
GCCAACGAAATCAACACTCTGGAAAACGTAGTTAACCAACAAAAAATTCCTCAAGTCACAAGAGTCGTCACAGAACA
ACAAGGTCGCAAGGTGCTAATTGACGGACATTGGGTGATAGACTTTGCTTCTTGCAACTATTTAGGTCTTGACTTGC
ATCCAAAAGTTAAGGAAGCAATTCCACCAGCTTTGGATAAATGGGGCACACATCCAAGCTGGACTCGGCTTGTTGCT
TCCCCAGCAATTTATGAGGAATTGGAGGAAGAATTGTCCAAACTTTTAGGCGTACCTGATGTTTTAGTATTTCCAGC
TGTAACACTGCTTCAGATAGGAATTTTACCACTATTAACTGGGAATAATGGTGTCATCTTTGGTGACATAGCTGCAC
ATCGTTGTATTTATGAAGCGTGCTGTCTGGCTCAGCACAAAGGAGCCCAGTTCATCCAATATCGACATAATGATTTG
AACGATTTAGCCGAAAAACTAGCAAAATATCCGCCTGAACAAGTAAAGATTATTGTCATTGATGGCGTGTATTCCAT
GTCGGCAGATTTTCCCGATCTGCCAGCTTACGTGCATCTGGCAAAAGAGTACAATGCCTTAATTTACATGGATGATG
CTCATGGTTTTGGCATTTTGGGCGAAAATCCCAGCAGCGATATGCCTTACGGTTACAAAGGAAACGGGATGGTGAAT
TATTTTGACCTGCGGTTTGCAGAGGATAATATCATCTATGTAGCTGGTTTGTCCAAAGCCTATTCTTCTTACGCAGC
ATTCTTAACTTGTGGCGATCGCCGGATCAAAACCAACTTCCGCAACGCTTGGACTGCCATATTTTCTGGTCCTTCTC
CTGTTGCGAGTTTGGCAAGTGCCTTAGCCGGATTACAGGTGAATCGTCAGGAGGGGGAGCAGTTAAGAAAACAAATT
TATCACCTAACTCACAAATTGGTTACACAAGCAAGAGCCATTGGATTCGAAGTGGATAACTATGGTTACGTTCCCAT
CGTAGGCGTGTTAGTGGGAGATGCTCAACACATGATTGATGTGTGTCAACTCCTTTGGGAATATGGTATTTTAATTA
CTCCTGCTATTTTTCCAATCGTACCTTTAAATAAAAGTGCTTTAAGGTTTTCGATTACAGCCGCCAATACCGAAGAG
GAGATAGACCAAGCAATTAAATCTCTCAAAGCAGTTTGGGATTTGCTACAAAAAAGGAAAGCTTTGCCTTGTAAGCA
GGAGGAAAACATACTCAAGCATgcggccgcactcgagcaccaccaccaccaccactga

SxtA(WT) amino acid sequence

(linkers and 6x-His tags are shown in red)

MGSSHHHHHHSSGLVPRGSHMLQKINRYTHGFVAVPVILACREKGVFELLADESPLSLNQMVEHLGANSGHFQVALR
MLESLHWLSRNKELKYSLTAEAAIHNKISEDILQLYNLPIQSYLEGKQGNLLGRWIERSCQLWNLDNPLMADFLDGL
LVIPLLLALHKHNLLADSEDKPLLSSLSSTVQEELGKLFLHLGWADLTAGRLTITELGRFMGERALNTAIVASYTPM
LSRIHDVLFGNCLSVFQRDASGHERHIDRTLNVIGSGFQHQKYFADLEESILSVFNQLPLEEQPKYITDMGCGDGTL
LKRVWETIQFKSARGKALEQYPLRLIGVDYNEASLKATTRTLASLPHLVLQGDIGNPEQMVRSLEAHGIHDPENILH
IRSFLDHDRLFIPPQKRNELKERAHLPYQSVCVDDQGELIPPHVMVQSLVEHLERWSQVVNKHGLMILEVHCLEPRV
VYQFLDKSENLHFDAFQGFSQQYLVEAEVFLMSAAQVGLFPKLELSKRYPKTFPFTRITLNYFEKRPYKISHAYLSD
LPALVDLEVKCWPENLRASTHEIRRRLELNPQGNLVLIIEDQIIGAIYSQTITSTEALENVKYAQVPTLHTPQGSVI
QLLALNILPEFQARGLGNELRDFMLYYCTLKGGIESVVGVTRCRNYVNYSQMPMMEYLKLHNEQRQLLDPIVGFHVS
GGAEIRGIIANYRPEDTDNLGMGILIEYNLRDSALHSPGDRKGPYINSAIGSLVPKATSATKENKTVADLVKECILK
VMGSQRQAAYAPQQKLLDMGLDSLDLLELQTLLEERLGINLSGTFFLQKNTPTAIITYFQNQVVQEKQSDLAPPVDS
ANEINTLENVVNQQKIPQVTRVVTEQQGRKVLIDGHWVIDFASCNYLGLDLHPKVKEAIPPALDKWGTHPSWTRLVA
SPAIYEELEEELSKLLGVPDVLVFPAVTLLQIGILPLLTGNNGVIFGDIAAHRCIYEACCLAQHKGAQFIQYRHNDL
NDLAEKLAKYPPEQVKIIVIDGVYSMSADFPDLPAYVHLAKEYNALIYMDDAHGFGILGENPSSDMPYGYKGNGMVN
YFDLRFAEDNIIYVAGLSKAYSSYAAFLTCGDRRIKTNFRNAWTAIFSGPSPVASLASALAGLQVNRQEGEQLRKQI
YHLTHKLVTQARAIGFEVDNYGYVPIVGVLVGDAQHMIDVCQLLWEYGILITPAIFPIVPLNKSALRFSITAANTEE
EIDQAIKSLKAVWDLLQKRKALPCKQEENILKHAAALEHHHHHH

C. Cloning and Mutagenesis

Domain boundaries used

Methyltransferase (MT): Met1 – Tyr504

GCN5-related *N*-acetyltransferase (GNAT): Phe505 – Ser721

Acyl carrier protein (ACP): Ala722 – Gln815

8-amino-7-oxononanoate synthase (AONS): Glu816 – His1245

Table 2.S1. Primers used to generate constructs in this Chapter

Name	Sequence
Primers for SxtA(S773A) by site-directed mutagenesis of whole plasmids	
SxtA S773A forward	5′-AAACTGCTGGATATGGGATTAGATGCGTTAGATTTATTAGAACTGCAAACG-3′
SxtA S773A reverse	5′-CGTTTGCAGTTCTAATAAATCTAACGCATCTAATCCCATATCCAGCAGTTT-3′
Primers for mutagenesis by splicing by overlap extension (SOE)	
SxtA H372A SOE forward	5′-AATATCCTGCACATCCGTTCGTTCCTCGACGCGGATCGTCTCTTTATT-3'
SxtA H372 SOE reverse	5′-GAACGGATGTGCAGGATATT-3′
SxtA T637V SOE forward	5′-ATTGAGAGCGTGGTGGGTGTAGTTCGCTGTCGAAATTATG-3′
SxtA T637 SOE reverse	5′-ACGCTCTCAATGCCGCCTTT-3′
SxtA H671A SOE forward	5′-CGACAGCTTCTGGATCCAATTGTGGGTTTCGCGGTGTCGGGAGGAGCC-3′
SxtA H671 SOE reverse	5′-ATTGGATCCAGAAGCTGTCG-3′
SxtA K1077L SOE forward	5′-GATAATATCATCTATGTAGCTGGTTTGTCCCTGGCCTATTCTTCTTAC-3′
SxtA K1077 SOE reverse	5′-CAAACCAGCTACATAGATGATATTATC-3′
Primers for ligation-independent cloning (LIC)	
SxtA MT M1 LIC forward	5′-**TACTTCCAATCCAATGCA**ATGTTACAAAAGATTAATCG-3′
SxtA GNAT F505 LIC forward	5′-**TACTTCCAATCCAATGCAATG**TTCGAGAAAAGACCTTACAA-3′
SxtA GNAT S721 LIC reverse	5′-**TTATCCACTTCCAATGCTA**TGAGTTAATATACGGTCCTT-3′
SxtA ACP A722 LIC forward	5′-**TACTTCCAATCCAATGCGATG**GCAATTGGTTCATTGGTAC-3′
SxtA ACP Q815 LIC reverse	5′-**TTATCCACTTCCAATGTTA**TTGTACCACTTGGTTTTGGAAA-3′
SxtA AONS H1245 LIC reverse	5′-**TACTTCCAATCCAATGCAATG**CAAGAGAAACAATCTGATCT-3′
SxtA AONS H1245 6H LIC reverse	5′-**TTATCCACTTCCAATGTTA**GTGGTGGTGGTGGTGGTGCTCGAG-3′
EC_FabD LIC forward	5′-**TACTTCCAATCCAATGCA**ATGACCCAATTTGCGTTTGTATTC-3′
EC_FabD LIC reverse	5′-**TTATCCACTTCCAATGTTA**CAGTTCCAAAGCGGCAGCC-3′

Overhangs for ligation-independent cloning are shown in boldface, mutations underlined

ACP knockout SxtA(S773A): pET28b-*sxtA*(S773A) was generated by site-directed mutagenesis on pET28b-*sxtA*(WT). 50 µL PCR reaction mixtures contained 10 µL Phusion HF buffer, 2 ng/µL WT parent plasmid, 2 µM each of the S773A forward and reverse primers, 200 µM each of dNTPs, 0.04 U/µL Phusion HF and 6% (v/v) DMSO. Amplification was accomplished with the following PCR procedure: 95 °C for 2:00, (95 °C 0:30, 60 °C 1:00, 68 °C 6:00) for 18 cycles, with a final extension of 68 °C for 15:00. This was followed by a 10 µL digestion containing 1 µL NEB CutSmart buffer, 8 µL of the PCR mixture and 20 units of DpnI. The reaction mixture was incubated at 37 °C for 3 h and transformed into chemically competent DH5α cells.

SxtA(H372A), SxtA(T637V), SxtA(H671A), and SxtA(K1077L): The remaining mutants were generated through splicing by overlap extension (SOE).[55] 50 µL PCR reaction mixtures for each fragment included 10 µL Phusion HF buffer, 2 ng/µL WT plasmid, 2 µM each of forward and reverse primers, 200 µM each of dNTPs, and 0.08 U/µL Phusion HF. *sxtA* fragments were amplified according to the following PCR procedure: 95 °C 2:00, (95 °C 0:30, 45-70 °C 1:00, 72 °C 0:45/kb) for 30 cycles, with a final extension of 72 °C for 15:00. C-terminal fragments included an additional 6x-His tag from pET28b-*sxtA*(WT). The PCR mixtures were purified with the QIAquick PCR Purification kit according to the manufacturer's instructions. The complete *sxtA* insert was assembled in 10 µL Phusion HF buffer, approx. 1:1 ratio of N- and C-terminal fragments, 200 µM each of dNTPs and 0.08 U/µL Phusion HF polymerase (50 µL total) according to the following PCR procedure: 95 °C 2:00, (95 °C 0:30, 45-65 °C 1:00, 72 °C 0:45/kb) for 15 cycles, 72 °C 15:00. Forward and reverse LIC primers (to 2 µM each), 200 µM each of additional dNTPs and 0.08 U/µL Phusion HF were then added, and the reaction was subjected to further amplification: 95 °C 2:00, (95 °C 0:30, 70 °C 1:00, 72 °C 4:00) for 30 cycles, 72 °C 15:00. The complete 4 kb PCR product was purified by gel extraction and subcloned into pMCSG7 using standard LIC protocols.[43]

MT-GNAT, GNAT, and ACP: Excised domains were amplified from pET28b-*sxtA*(WT) in 50 µL reactions containing 10 µL Phusion HF buffer, 2 ng/µL parent plasmid, 2 µM each of forward and reverse LIC primers, 200 µM each of dNTPs, 0.04 U/µL Phusion HF. Excised domains were amplified according to the following PCR procedure: 95 °C 2:00, (95 °C 0:30, 46-50 °C 0:30, 72 °C 0:45/kb) for 30 cycles, 72 °C 10:00. Inserts were purified by gel extraction and subcloned into pMCSG7 using standard LIC protocols.[43]

ACP(KO)-AONS: The excised didomain with the ACP knockout and no additional C-terminal 6x-His tag was amplified from pET28b-*sxtA*(S773A) in a 50 µL reaction containing 10 µL Phusion HF buffer, 2 ng/µL parent plasmid, 2 µM each of forward and reverse LIC primers, 200 µM each of dNTPs, 0.04 U/µL Phusion HF. The didomain was amplified according to the following PCR procedure: 95 °C 2:00, (95 °C 0:30, 55 °C 1:00, 72 °C 2:00) for 30 cycles, 72 °C 10:00. The insert was purified by gel extraction and subcloned into pMCSG7 using standard LIC protocols.[43]

***E. coli* FabD:** The synthetic gene was amplified in a 50 µL reaction containing 10 µL Phusion HF buffer, 2 ng/µL *fabD*, 2 µM each of forward and reverse LIC primers, 200 µM each of dNTPs, 0.04 U/µL Phusion HF. The didomain was amplified according to the following PCR procedure: 95 °C 2:00, (95 °C 0:30, 62 °C 0:45, 72 °C 1:00) for 25 cycles, 72 °C 10:00. The insert was purified by gel extraction and subcloned into pMCSG7 using standard LIC protocols.

All mutations were confirmed by Sanger sequencing at the University of Michigan DNA Sequencing Core.

D. Protein overexpression and purification

Overexpression of full SxtA modules (WT, H372A, T637V, H671A, and K1077L) and stringently purified *holo*-ACP: pET28b or pMCSG7 plasmids containing complete *sxtA* were transformed into chemically competent BAP1 *E. coli*. A single colony was picked to inoculate a 5 mL LB starter culture grown overnight at 37 °C, 200 rpm. The following day, 0.5 L TB media supplemented with kanamycin or ampicillin was inoculated with the starter culture and incubated

at 37 °C, 250 rpm until the OD_{600} reached 1.0. Cultures were equilibrated at 20 °C for 1 h prior to induction by addition of isopropyl-β-D-thiogalactopyranoside (IPTG, final concentration 200 μM). Cultures were incubated at 20 °C, 200 rpm for 18 h.

Overexpression of SxtA(S773A) and excised domains (MT-GNAT, GNAT, *apo*-ACP, ACP(KO)-AONS): pET28b or pMCSG7 plasmids containing the desired insert were transformed into pRARE-containing chemically competent BL21(DE3) *E. coli* cells. A single colony was picked to inoculate a 5 mL LB starter culture grown overnight at 37 °C, 200 rpm. The following day, 0.5 L TB media supplemented with spectinomycin and either kanamycin or ampicillin, was inoculated with the starter culture and incubated at 37 °C, 250 rpm until the OD_{600} reached 1.0. For *apo*-ACP, the media was also supplemented with a trace metals mixture (final concentrations: 50 μM $FeCl_3$, 20 μM $CaCl_2$, 10 μM $MnCl_2$, 10 μM $ZnCl_2$, 2 μM $CoCl_2$, 2 μM $CuCl_2$, 2 μM $NiCl_2$, 2 μM Na_2MoO_4, 2 μM Na_2SeO_3, 2 μM $B(OH)_3$). Cultures were equilibrated at 20 °C for 1 h. Expression was induced by addition of IPTG (final concentration 200 μM). Cultures were incubated at 20 °C, 200 rpm for 18 h.

Overexpression of *holo*-ACP, Sfp, *E. coli* FabD: pMCSG7-*sxtA*_ACP, pET24b-*sfp*, and pMCSG7-*ECfabD* were transformed into chemically competent BAP1 and BL21(DE3), respectively. A single colony was picked to inoculate a 10 mL LB starter culture grown overnight at 37 °C, 200 rpm. The following day, 1 L LB cultures of *holo*-ACP and Sfp were supplemented with ampicillin or kanamycin, respectively, inoculated with the starter cultures, and incubated at 37 °C, 200 rpm until the OD_{600} reached 0.6. The culture was cooled to rt and expression was

induced by addition of IPTG (final concentration 100 μM). Cultures were incubated at 18 °C, 200 rpm for 18 h.

FPLC purification of SxtA(WT) and SxtA(S773A): Cell pellets were resuspended in 4 mL of lysis buffer (50 mM HEPES, 300 mM NaCl, 10 mM imidazole, 1 mM pyridoxal phosphate (PLP), 10% (v/v) glycerol, pH 8.0) per gram of wet cell mass. Protease inhibitors pepstatin A (1 μg/mL), benzamidine hydrochloride (15 μg/mL), and phenylmethylsulfonyl fluoride (PMSF, 1 mM) were added to the resuspension mixture before incubation at 4 °C for 20 min with gentle shaking. The resuspended mixture was lysed by homogenization (Avestin EmulsiFlex-C3, 15,000 psi, 4 passes). Insoluble material was removed by centrifugation twice (20,000 x *g* for 20 min and 30,000 x *g* for 30 min, both at 4 °C). The clarified lysate was sterile filtered (0.45 μm syringe filters, CELLTREAT Scientific Products) and loaded onto a 5 mL HisTrap column. FPLC buffers: Buffer A = 50 mM HEPES, 300 mM NaCl, 10 mM imidazole, 10% (v/v) glycerol, pH 8.0; buffer B = 50 mM HEPES, 300 mM NaCl, 300 mM imidazole, 10% (v/v) glycerol, pH 8.0. The column was washed with 6 column volumes (CV) of 10% B followed by a linear gradient of 10-25% B over 2 CV. The desired protein was eluted with 25% B over 4 CV followed by a linear gradient of 25-75% B over 7 CV. Fractions containing the protein of interest were pooled, concentrated using 50 kDa centrifugal cutoff filters, and exchanged into a storage buffer consisting of 50 mM HEPES, 200 mM NaCl, and 10% (v/v) glycerol) at pH 7.4 using a PD-10 column. The desalted protein was concentrated further using a 50 kDa centrifugal cutoff filter, aliquoted, flash frozen in liquid nitrogen, and stored at -80 °C.

Stringent purification of *holo*-ACP: Cell pellets were resuspended in 4 mL of lysis buffer (50 mM HEPES, 300 mM NaCl, 10 mM imidazole, and 10% (v/v) glycerol at pH 8.0) per gram of wet

cell mass. Cells were lysed by sonication (3 s on, 6 s off, 5 min total). Insoluble material was removed by centrifugation (30,000 x *g* for 30 min at 4 °C). The clarified lysate was sterile filtered (0.45 μm syringe filters, CELLTREAT Scientific Products) and loaded onto a 5 mL HisTrap column (GE Healthcare). Nickel affinity FPLC buffers: buffer A = lysis buffer; buffer B = 50 mM HEPES, 300 mM NaCl, 300 mM imidazole, 10% (v/v) glyercol, pH 8.0. The column was washed with 3 CV of 5 % B, followed by a linear gradient of 5-100% B over 5 CV. Fractions containing ACP were pooled and concentrated using 10 kDa centrifugal cutoff filters. The concentrated protein was diluted with anion exchange binding buffer (50 mM HEPES, 10% (v/v) glycerol, pH 8.0) and loaded onto a 5 mL HiTrap Q HP column (GE Healthcare). Anion exchange FPLC buffers: buffer C = binding buffer; buffer D = 50 mM HEPES, 1 M NaCl, 10% (v/v) glycerol, pH 8.0. The column was washed with 5 CV of buffer C. The desired protein was eluted with a linear gradient 0-100% D over 10 CV. Fractions containing ACP were pooled, concentrated using 10 kDa centrifugal cutoff filters and further purified by size exclusion chromatography (GE Healthcare Superdex 75 Increase 10/300 GL) into the storage buffer (50 mM HEPES, 200 mM NaCl, pH 7.4). Fractions containing ACP were pooled and analyzed by mass spectrometry (see section "Enzymatic reactions and data" below for details). The purified ACP was found to be 72% *holo*, 28% *apo*. The remaining apo-ACP was enzymatically converted to *holo*-ACP in a 1 mL reaction containing 2 mM CoASH, 500 μM ACP, 25 μM Svp (from *Streptomyces verticillus*, kindly donated by the group of Janet Smith)[56] in 50 mM Tris, 150 mM NaCl, 20 mM $MgCl_2$, 10% (v/v) glycerol at pH 7.4 at 30° C for 2 h. *Holo*-ACP was separated by exclusion chromatography again into the storage buffer. Fractions containing pure ACP were pooled, concentrated using 10 kDa centrifugal cutoff filters, aliquoted, flash frozen in liquid nitrogen, and stored at -80 °C.

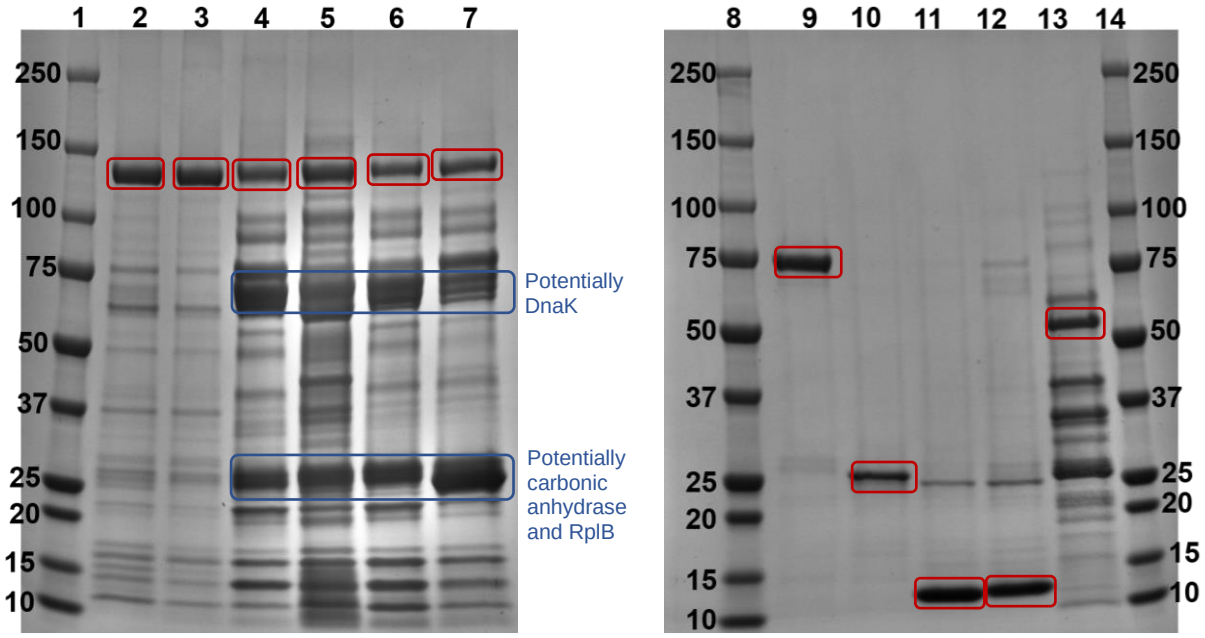
Purification of all other SxtA constructs: Cell pellets were resuspended in 4 mL of lysis buffer (50 mM HEPES, 300 mM NaCl, 10 mM imidazole, and 10% (v/v) glycerol at pH 8.0) per gram of wet cell mass. For AONS-containing constructs (except SxtA(K1077L)), 1 mM PLP was added to the lysis buffer. Protease inhibitors pepstatin A (1 μg/mL), benzamidine hydrochloride (15 μg/mL), and PMSF (1 mM) were added to the resuspension mixture before incubation at 4 °C for 20 min with gentle shaking. Cells were lysed by sonication (3 s on, 6 s off, 5 min total). Insoluble material was removed by centrifugation (30,000 x *g* for 30 min at 4 °C). The clarified lysate was incubated with gentle shaking along with 1-2 mL of nickel-NTA resin for 1 h at 4 °C and poured into a column. The resin-bound protein was washed with 25 mL wash buffer (50 mM HEPES, 300 mM NaCl, 25 mM imidazole, 10% (v/v) glycerol, pH 8.0). The desired protein was eluted with 6 mL elution buffer (50 mM HEPES, 300 mM NaCl, 300 mM imidazole, 10% (v/v) glyercol, pH 8.0). The eluted protein was concentrated using 3 – 50 kDa centrifugal cutoff filters and exchanged into storage buffer (50 mM HEPES, 200 mM NaCl, 10% (v/v) glycerol, pH 7.4) using a PD-10 column. The desalted protein was concentrated further using 3 – 50 kDa centrifugal cutoff filters, aliquoted, flash frozen in liquid nitrogen and stored at -80 °C.

Purification of Sfp and *E. coli* FabD: Cell pellets were resuspended in 4 mL of lysis buffer (50 mM Tris, 300 mM NaCl, 10 mM imidazole, and 10% (v/v) glycerol, pH 7.4) per gram of wet cell mass. Cells were lysed by sonication (3 s on, 3 s off, 3 min total). Insoluble material was removed by centrifugation (30,000 x *g* for 30 min at 4 °C). The clarified lysate was incubated with 1-2 mL of nickel-NTA resin for 1 h at 4 °C and poured into a column. The resin-bound protein was washed with 25 mL wash buffer (50 mM Tris, 300 mM NaCl, 20 mM imidazole, 10% (v/v) glycerol, pH 7.4). The desired protein was eluted with 6 mL elution buffer (50 mM Tris, 300 mM NaCl, 300

mM imidazole, 10% (v/v) glycerol, pH 7.4). The eluted protein was concentrated using a 10 kDa centrifugal cutoff filter and exchanged into the storage buffer (50 mM Tris, 10% (v/v) glycerol, pH 7.4) using a PD-10 column. The desalted protein was concentrated further using a 10 kDa centrifugal cutoff filter, aliquoted, flash frozen in liquid nitrogen, and stored at -80 °C.

Table 2.S2. SxtA protein constructs and calculated molecular weight with intact initial Met

Construct name	Plasmid	Expression line	Calculated protein mass (Da)
SxtA(WT)	pET28b	BAP1	143,548.8
SxtA(S773A)	pET28b	BL21(DE3) pRARE	143,532.8
SxtA(H372A)	pMCSG7	BAP1	144,057.3
SxtA(T637V)	pMCSG7	BAP1	144,121.4
SxtA(H671A)	pMCSG7	BAP1	144,057.3
SxtA(K1077L)	pMCSG7	BAP1	144,108.4
SxtA MT-GNAT	pMCSG7	BL21(DE3) pRARE	84,167.6
SxtA GNAT	pMCSG7	BL21(DE3) pRARE	27,455.3
SxtA *apo*-ACP	pMCSG7	BL21(DE3) pRARE	13,193.1
SxtA *holo*-ACP	pMCSG7	BAP1	13,533.2
SxtA ACP(KO)-AONS	pMCSG7	BL21(DE3) pRARE	60,974.0
	pET24b	BL21(DE3)	28,544.9



1: ladder (kDa)
2: SxtA(WT)
3: SxtA(S773A)
4: SxtA(H372A)
5: SxtA(T637V)
6: SxtA(H671A)
7: SxtA(K1077L)

8: ladder (kDa)
9: SxtA MT-GNAT
10: SxtA GNAT
11: SxtA *apo*-ACP
12: SxtA *holo*-ACP
13: SxtA ACP(KO)-AONS
14: ladder (kDa)

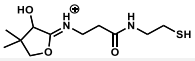
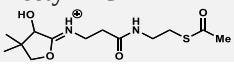
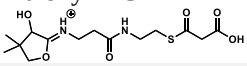
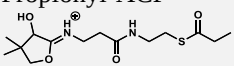
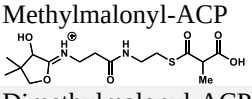
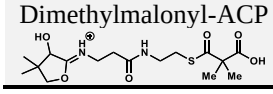
Figure 2.S1. Denatured SDS-PAGE protein gels.
Proteins of interest are highlighted in red. Common co-purifying proteins (blue), potentially carbonic anhydrase (25.1 kDa), RplB (29.9 kDa) and DnaK (69.1 kDa).[57] Further purification diminished activity.

III. Enzymatic reactions

A. General information

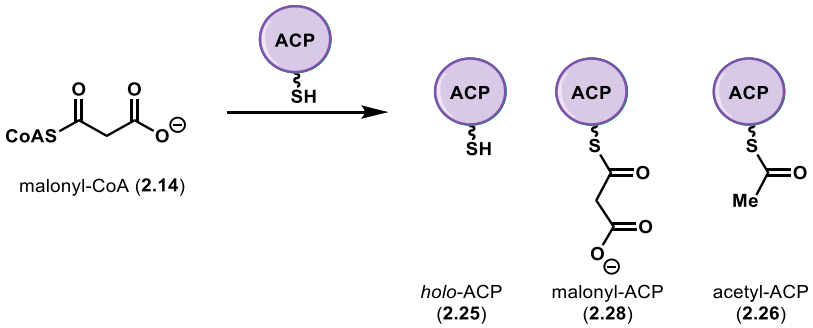
CoASH, butyryl-, isobutyryl-, isovaleryl-, octanoyl- and benzoyl-CoA were purchased from Sigma-Aldrich. Acetyl-, malonyl-, propionyl-, methylmalonyl-, and hexanoyl-CoA were obtained from CoALA Biosciences. *S*-adenosylmethionine was purchased as the tosyl salt from Carbosynth and prepared as a 50 mM aqueous solution immediately before use. Electrospray liquid chromatography-mass spectrometry (LC-MS) analysis was performed on an Agilent G6545A quadrupole-time of flight mass spectrometer in positive mode with an Agilent 1290 UPLC system. Solvent A = water with 0.1% formic acid. Solvent B = 95% acetonitrile, 5% water and 0.1% formic acid. MS data were analyzed using the Agilent Qualitative Mass Hunter software. Intact proteins were deconvoluted with the maximum entropy algorithm.

Table 2.S3. Acyl-ACP masses

Species	Calculated intact protein mass (Da)	Observed mass (Da)	Calculated Ppant ejection[44] fragment (*m/z*)	Observed Ppant ejection mass (*m/z*)
Apo-ACP	13,193.1	13,193.2	N/A	N/A
Holo-ACP 	13,533.2	13,533.2	261.13	261.13
Acetyl-ACP 	13,575.2	13,575.4	303.14	303.14
Malonyl-ACP 	13,619.2	13,619.4	347.13	347.13
Propionyl-ACP 	13,589.2	13,589.5	317.15	317.15
Methylmalonyl-ACP 	13,633.2	13,633.1	361.14	361.14
Dimethylmalonyl-ACP 	13,647.2	ND	375.16	ND

ND: not detected

B. Enzymatic reactions and data



Loading of *holo*-ACP (Figure 2.8): 500 μM acyl-CoA, 100 μM *holo*-ACP, optional catalysts, in 50 mM HEPES pH 7.4, 150 mM NaCl, 1 mM $MgCl_2$, 0.5 mM $MnCl_2$ were combined to a final volume of 60 μL, adding acyl-CoA last. Catalysts: 0 – 15 μM GNAT construct, or 50 nM FabD from *Moorea producens* or *E. coli*. Reactions were incubated at 30 °C for 2 h. At 1, 2, 10, 30, 60 and 120 min, 10 μL aliquots were removed and quenched by diluting 12.5x with 1% formic acid in water. Samples were analyzed using a Ppant ejection assay of intact proteins.[44] LC-MS analysis: column = Phenomenex Aeris 3.6 μm WIDEPORE C4 2.1 x 50 mm; method = 5% B at 0.5 mL/min for 2 min, followed by a linear gradient to 100% B over 4 min, 100% B for 2 min, followed by a 0.1 min linear gradient to 5% B and 1.9 min equilibration at 5% B (total time 10 min). t_R = 4.1 min.

For malonyl-CoA and methylmalonyl-CoA, the relative abundance of ejected Ppant ions for *holo*, and decarboxylated (acetyl or propionyl) plus carboxylated (malonyl or methylmalonyl) ACP were used to calculate the percentage of loaded ACP.

For all other acyl-CoAs, the percentage of loaded ACP was calculated from the relative abundance of *holo*- and acyl-ACP ejected Ppant ions.

All loading reactions were run in duplicate.

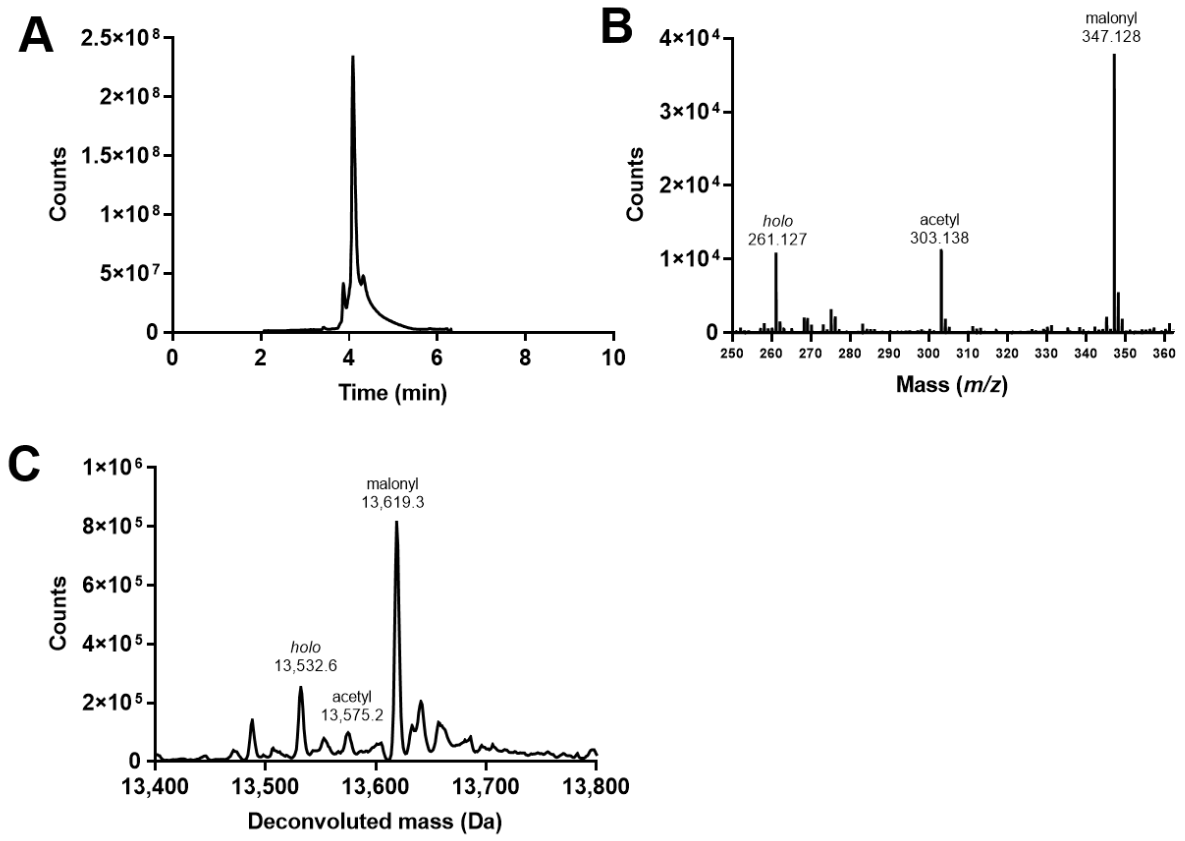


Figure 2.S2. Representative LC-MS analysis of *holo*-ACP loading with malonyl-CoA.
(A) Total ion chromatogram; (B) mass spectrum of ejected Ppant ions at 4.1 min; (C) deconvolution of the peak at 4.1 min. Minor species resulting from the loss of the initial N-terminal methionine were also observed (-131 Da).

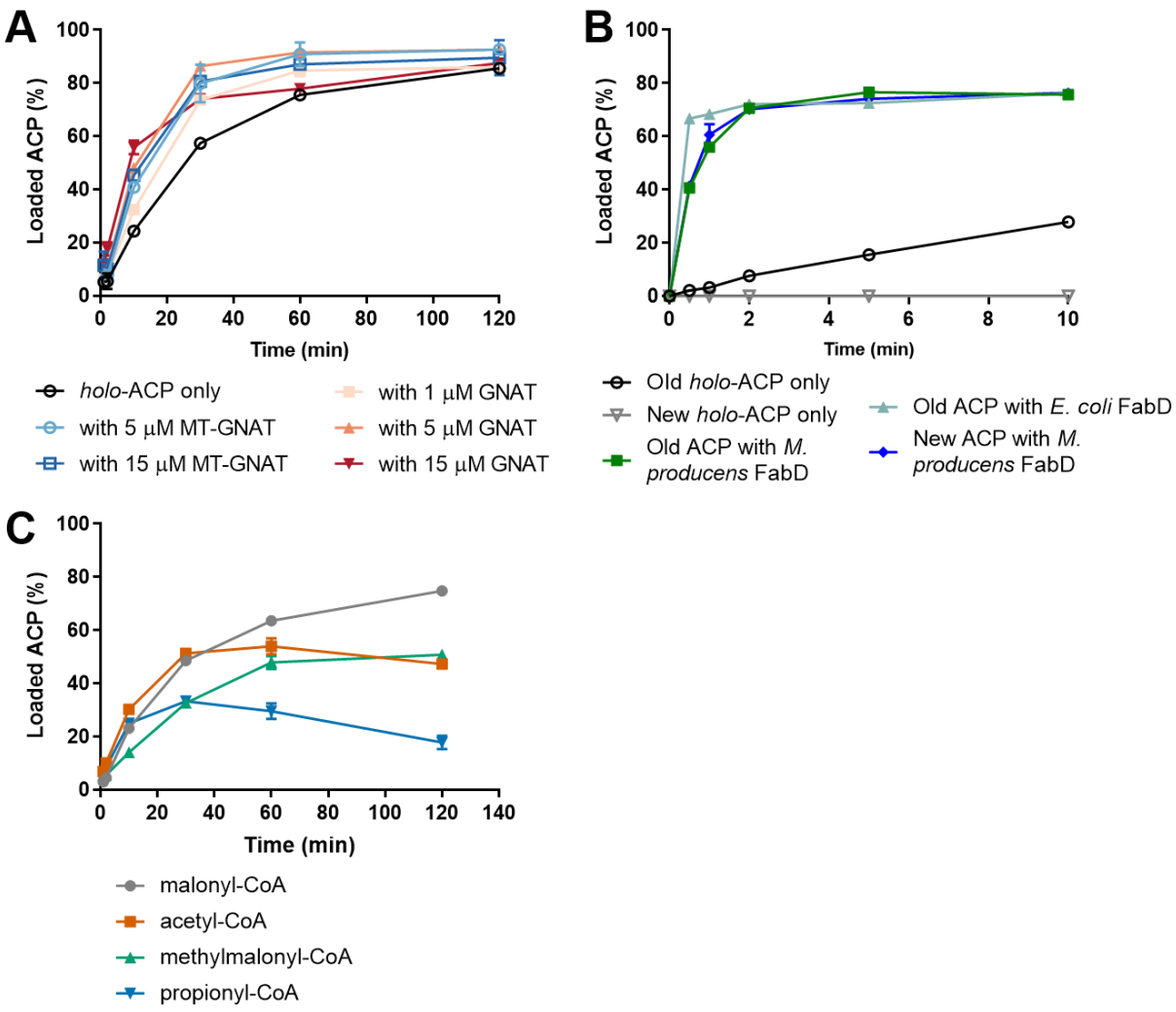
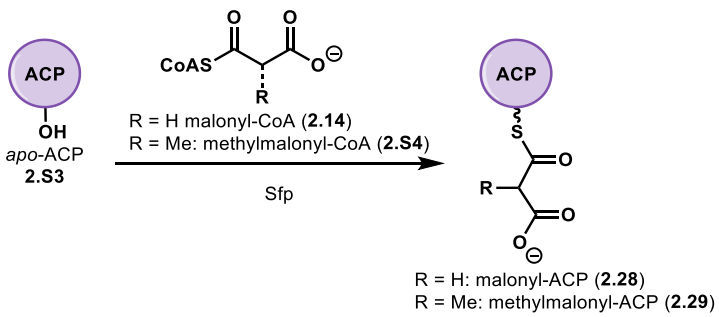


Figure 2.S3. SxtA ACP is likely loaded by an acyltransferase favoring malonyl-CoA.
(A) Loading of *holo*-ACP with malonyl-CoA, GNAT constructs added. (B) Loading of *holo*-ACP with malonyl-CoA, in the presence of FabD. (C) Loading of *holo*-ACP with additional acyl-CoAs (no GNAT constructs added).

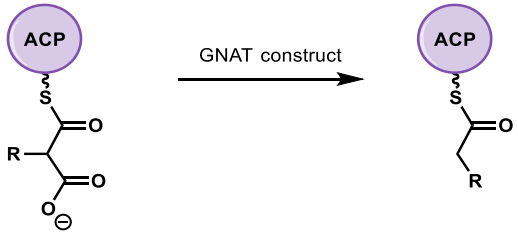


Preparation of malonyl-ACP and methylmalonyl-ACP (for Figures 2.9 and 2.10): 800 μM acyl-CoA, 180 μM *apo*-ACP, 20 μM Sfp in 50 mM Tris, 150 mM NaCl, 20 mM $MgCl_2$, 10% (v/v) glycerol at pH 7.4 were combined to a total volume of 2 mL. Reaction mixtures were incubated at 30 °C for 4 h. Acyl-ACP was separated by size exclusion chromatography (GE Healthcare HiPrep 16/60 Sephacryl S-200 High Resolution) using storage buffer (50 mM HEPES, 200 mM NaCl, pH 7.4) and stored at -80 °C. The modification was verified using a Ppant ejection assay of intact proteins.[44] LC-MS analysis: column = Phenomenex Aeris 3.6 μm WIDEPORE C4 2.1 x 50 mm; method = 5% B at 0.5 mL/min for 2 min, followed by a linear gradient to 100% B over 4 min, 100% B for 2 min, followed by a 0.1 min linear gradient to 5% B and 1.9 min equilibration at 5% B (total time 10 min). t_R = 4.1 min. The relative abundance of ejected Ppant ions for carboxylated ACP (malonyl or methylmalonyl) to decarboxylated ACP (acetyl or propionyl) were used to calculate the fraction of each species.

ACP acylation:

Malonyl-CoA: 81% malonyl-ACP, 19% acetyl-ACP

Methylmalonyl-CoA: 46% methylmalonyl-ACP, 54% propionyl-ACP



R = H: malonyl-ACP (**2.28**)
R = Me: methylmalonyl-ACP (**2.29**)

R = H: acetyl-ACP (**2.26**)
R = Me: propionyl-ACP (**2.27**)

Decarboxylation of malonyl- and methylmalonyl-ACP (Figure 2.10): 100 μM malonyl- or methylmalonyl-ACP and 0-15 μM GNAT construct in 50 mM HEPES pH 7.4, 150 mM NaCl, 1 mM $MgCl_2$, 0.5 mM $MnCl_2$ were combined to a final volume of 35 μL and incubated at 30 °C for 2 h. At 2, 10, 30, 60, 120 and 180 min, 5 μL aliquots were quenched by diluting 12.5x with 1% formic acid in water. Samples were analyzed using a Ppant ejection assay of intact proteins.[44] Decarboxylation reactions were run in duplicate. LC-MS analysis: column = Phenomenex Aeris 3.6 μm WIDEPORE C4 2.1 x 50 mm; method = 5% B at 0.5 mL/min for 2 min, followed by a linear gradient to 100% B over 4 min, 100% B for 2 min, followed by a 0.1 min linear gradient to 5% B and 1.9 min equilibration at 5% B (total time 10 min). t_R = 4.1 min.

The relative abundance of ejected Ppant ions for acetyl-/propionyl- and malonyl-/methylmalonyl-ACP were used to calculate the average formation of decarboxylated ACP at each time point. The background decarboxylation (i.e., negative control with no MT-GNAT or GNAT addition) was then subtracted from the reaction to obtain the data in Figure S5. The hydrolysis product *holo*-ACP was excluded from the decarboxylation analysis.

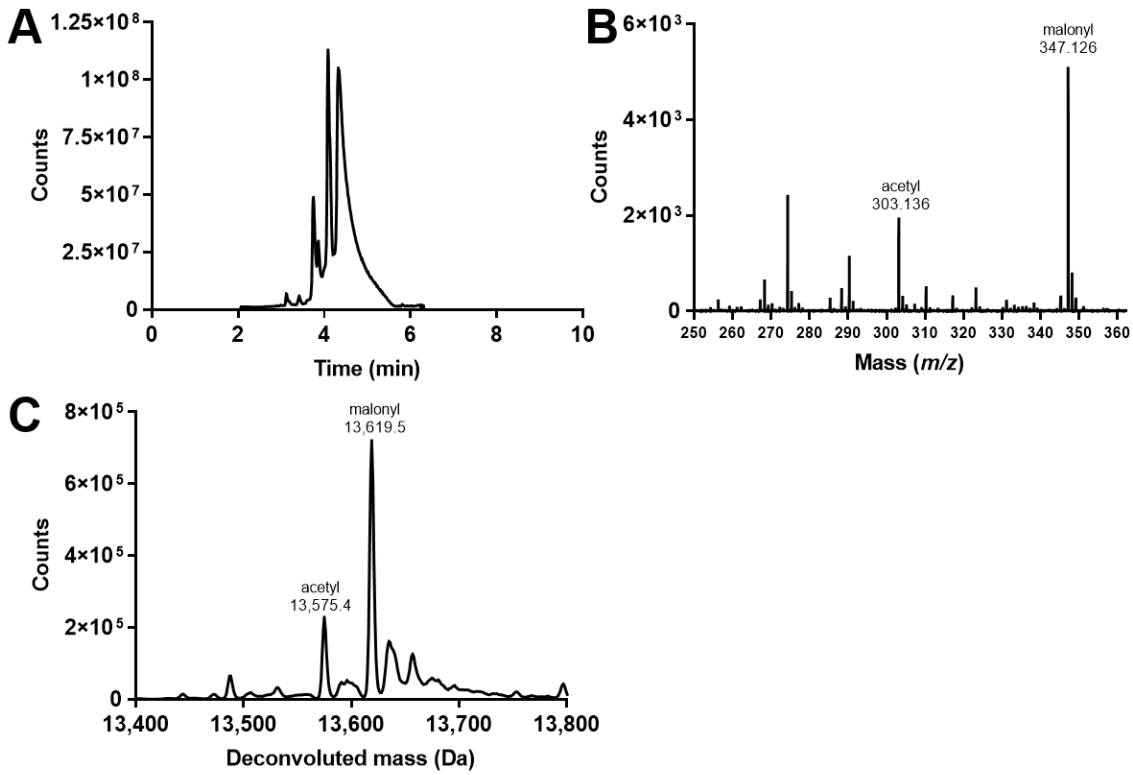


Figure 2.S4. Representative LC-MS analysis of malonyl-ACP decarboxylation to acetyl-ACP.

(A) Total ion chromatogram; (B) mass spectrum of ejected Ppant ions at 4.1 min; (C) deconvolution of the peak at 4.1 min. Minor species resulting from the loss of the initial N-terminal methionine were also observed (-131 Da).

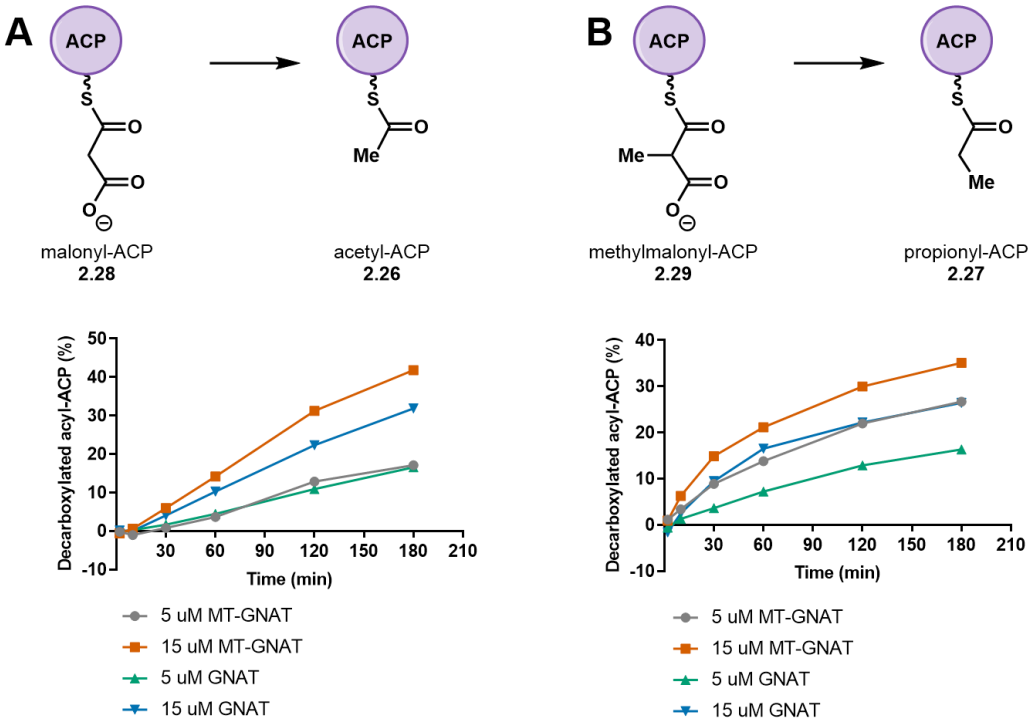


Figure 2.S5. Decarboxylation of carboxylated ACPs.
(A) Formation of acetyl-ACP from malonyl-ACP with background spontaneous decarboxylation subtracted. (B) Formation of propionyl-ACP from methylmalonyl-ACP with background spontaneous decarboxylation subtracted.

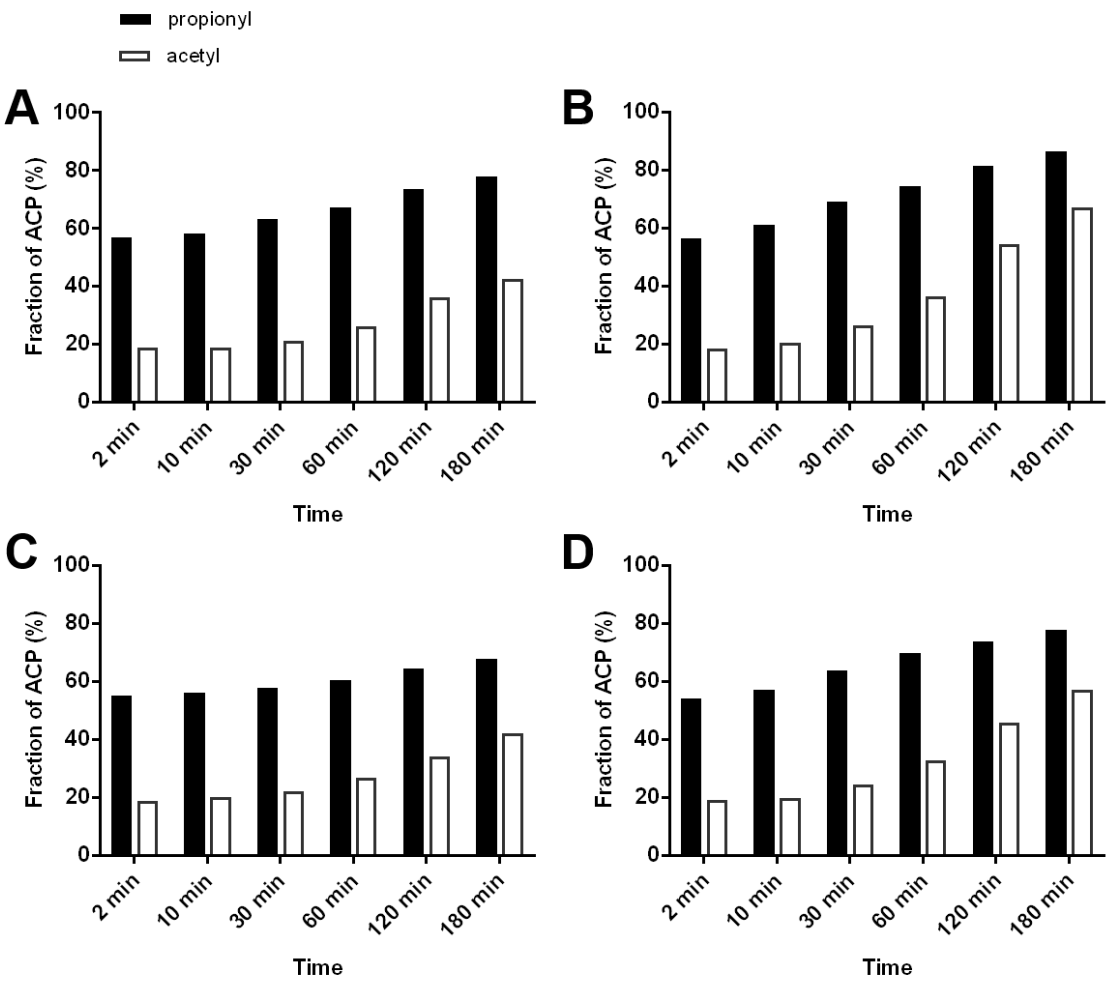
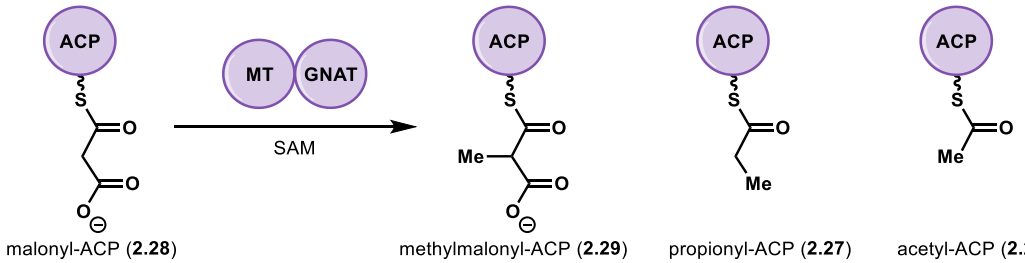


Figure 2.S6. Comparison of propionyl- and acetyl-ACP formation.
Black bars, from methylmalonyl-ACP; white bars, from malonyl-ACP; over 3 h with background spontaneous decarboxylation subtracted. (A) Acyl-ACPs with 5 µM MT-GNAT; (B) with 15 µM MT-GNAT; (C) with 5 µM GNAT; (D) with 15 µM GNAT.



Methylation of malonyl-ACP (Figure 2.9): 100 µM malonyl-ACP and 0-15 µM MT-GNAT in 50 mM HEPES pH 7.4, 150 mM NaCl, 1 mM $MgCl_2$, 0.5 mM $MnCl_2$, and 1 mM SAM were combined for a total volume of 35 µL and incubated at 30 °C for 3 h. At 15, 30, 45, 60, 120, and 180 min, 5 µL aliquots were removed and quenched by diluting 12.5x with 1% formic acid in water. Samples were analyzed using a Ppant ejection assay of intact proteins.[44] LC-MS analysis: column = Phenomenex Aeris 3.6 µm WIDEPORE C4 2.1 x 50 mm; method = 5% B at 0.5 mL/min for 2 min, followed by a linear gradient to 100% B over 4 min, 100% B for 2 min, followed by a 0.1 min linear gradient to 5% B and 1.9 min equilibration at 5% B (total time 10 min). t_R = 4.1 min. The relative abundance of ejected Ppant ions for acetyl-, malonyl-, propionyl- and methylmalonyl-ACP were used to calculate the fraction of each acyl-ACP species. They hydrolysis product *holo*-ACP was excluded from the methylation analysis. Methylation reactions were run in duplicate.

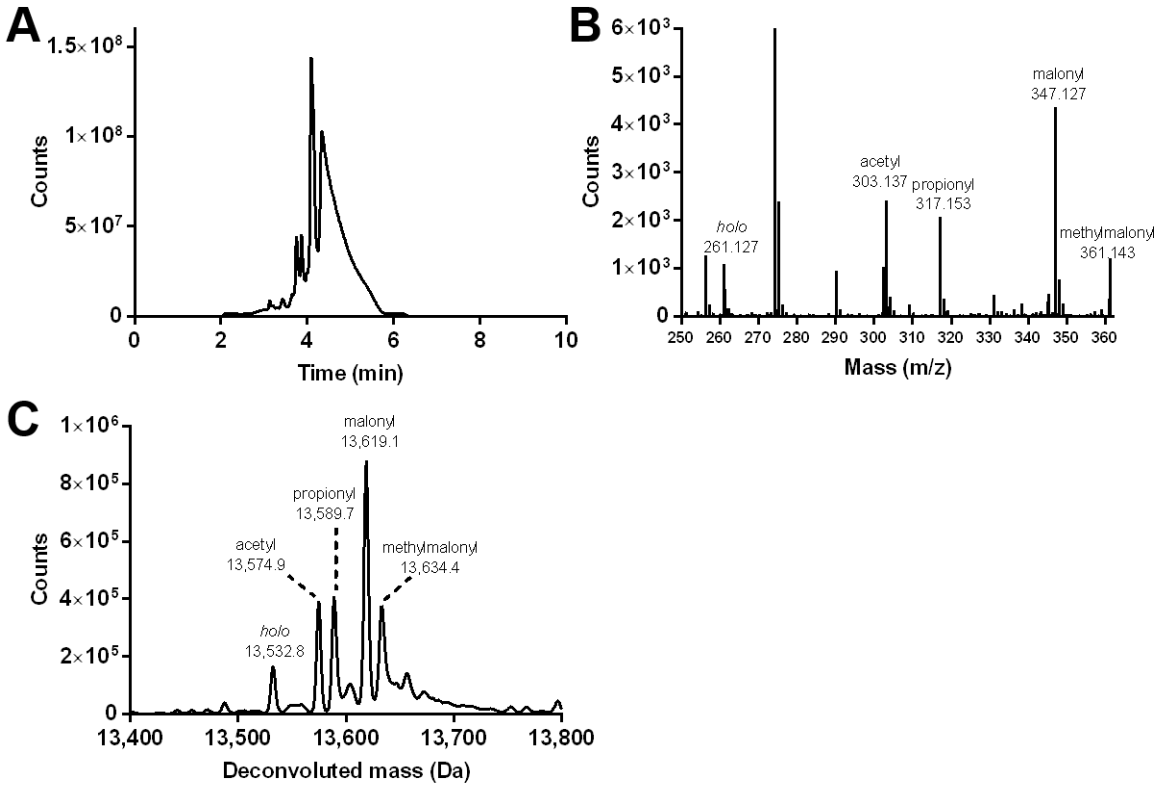


Figure 2.S7. Representative LC-MS analysis of malonyl-ACP methylation.
(A) Total ion chromatogram; (B) mass spectrum of ejected Ppant ions at 4.1 min; (C) deconvolution of the peak at 4.1 min. Minor species resulting from the loss of the initial N-terminal methionine were also observed (-131 Da).

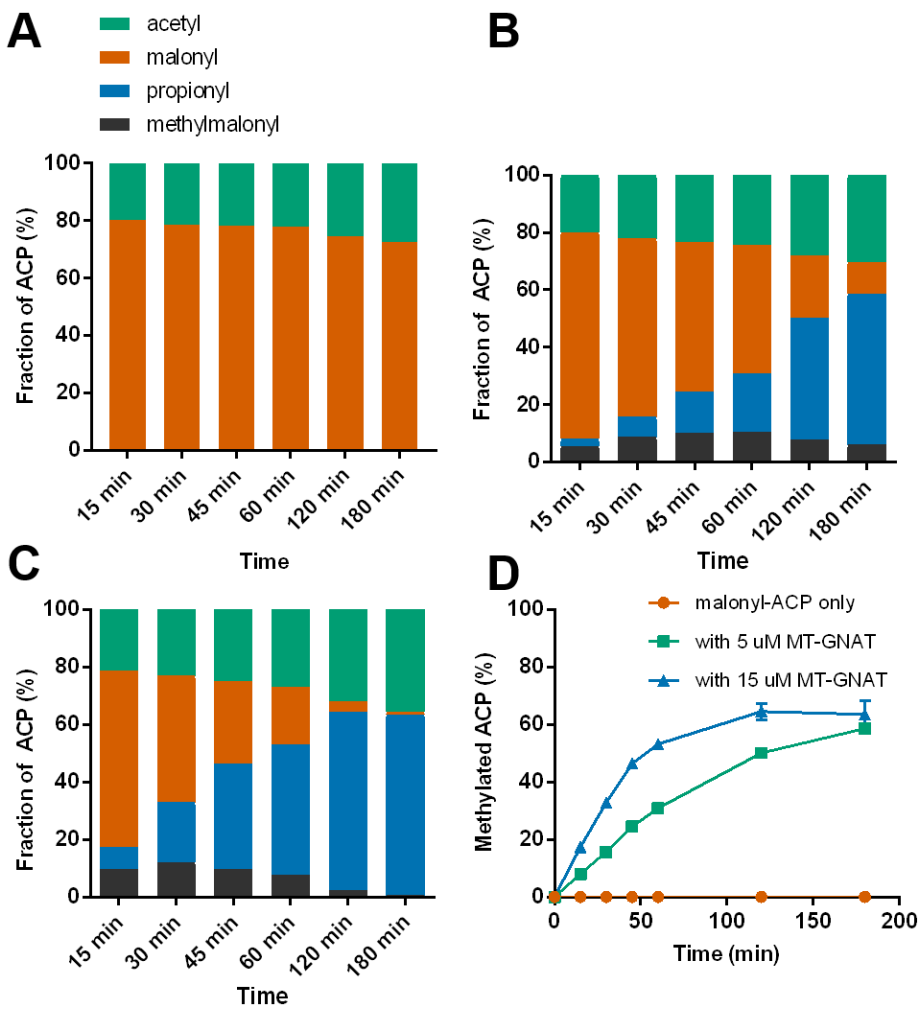
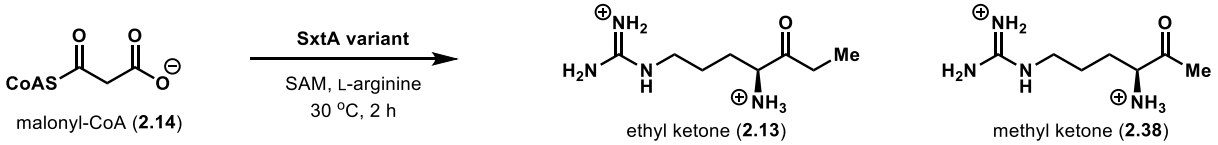


Figure 2.S8. Total methylation of malonyl-ACP.
(A) Malonyl-ACP without MT-GNAT; (B) malonyl-ACP with 5 μM MT-GNAT; (C) malonyl-ACP with 15 μM MT-GNAT; (D) total fraction of methylated ACP (methylmalonyl- and propionyl-ACP).



Full SxtA module reactions (Tables 2.2 and 2.5): 250 μM acyl-CoA, 1 mM SAM, 2 mM Arg, 10 μM SxtA (with 50 μM *holo*-ACP where applicable) in 50 mM HEPES pH 7.4, 150 mM NaCl, 1 mM $MgCl_2$, 0.5 mM $MnCl_2$ (total volume 25 μL) were combined and incubated at 30 °C for 2 h. Reactions were quenched by diluting 5x with acetonitrile. Precipitate was pelleted at 12,000 x *g* for 20 min. The supernatant was diluted 20x with acetonitrile, 1% formic acid, and aqueous 10 μg/mL ^{15}N Arg·HCl. The final ratios (v/v) for analysis were 5% quenched reaction, 2.5% of aqueous 10 μg/mL ^{15}N Arg·HCl (Cambridge Isotope Laboratories), 46% acetonitrile, 46.5% of 1% formic acid in water (e.g., 10 μL of quenched reaction, 5 μL of 10 μg/mL ^{15}N Arg·HCl, 92 μL acetonitrile, 93 μL of 1% formic acid in water for a total volume of 200 μL). Samples were analyzed by LC-MS: column = Waters Acquity 1.7 μm UPLC BEH Amide HILIC 2.1 x 100 mm column; method = 85% B at 0.3 mL/min for 2 min, followed by a linear gradient to 75% B at 0.4 mL/min over 3 min and a second linear gradient to 60% B over 0.5 min, 60% B for 1 min and then 6.5 min re-equilibration at 85% B at 0.3 mL/min (total time 15 min). t_R (**2.13**): 1.94 min; t_R (**2.38**): 2.35 min; t_R (^{15}N Arg): 4.0 min. Yields were calculated by comparing the areas under the curve of the extracted ion chromatograms to the ^{15}N Arg internal standard and normalized to the SxtA(WT) reaction with malonyl-CoA, SAM and Arg (Table 1, entry 3 or Table 2, entry 1). Reactions were run in triplicate.

Entry 8 was performed about a year later: 4 mM ATP, 2 mM methionine, 1.5 mg/mL *E. coli* MetK (provided by the laboratory of Jennifer Bridwell-Rabb at the University of Michigan) in 100 mM Tris pH 7.7, 200 mM KCl, 20 mM $MgCl_2$ (total volume 25 μL) were combined and incubated for 3 h at 37 °C. 0.5 mM $MnCl_2$, 0.1 mM PLP, 2 mM Arg, 500 μM malonyl-CoA and 50 nM of *E. coli* FabD were added to a final volume of 50 μL and incubated further for 2 h at 30 °C. The

reaction was quenched instead by diluting 5x with methanol, but the rest of the sample preparation and MS analysis was the same as above.

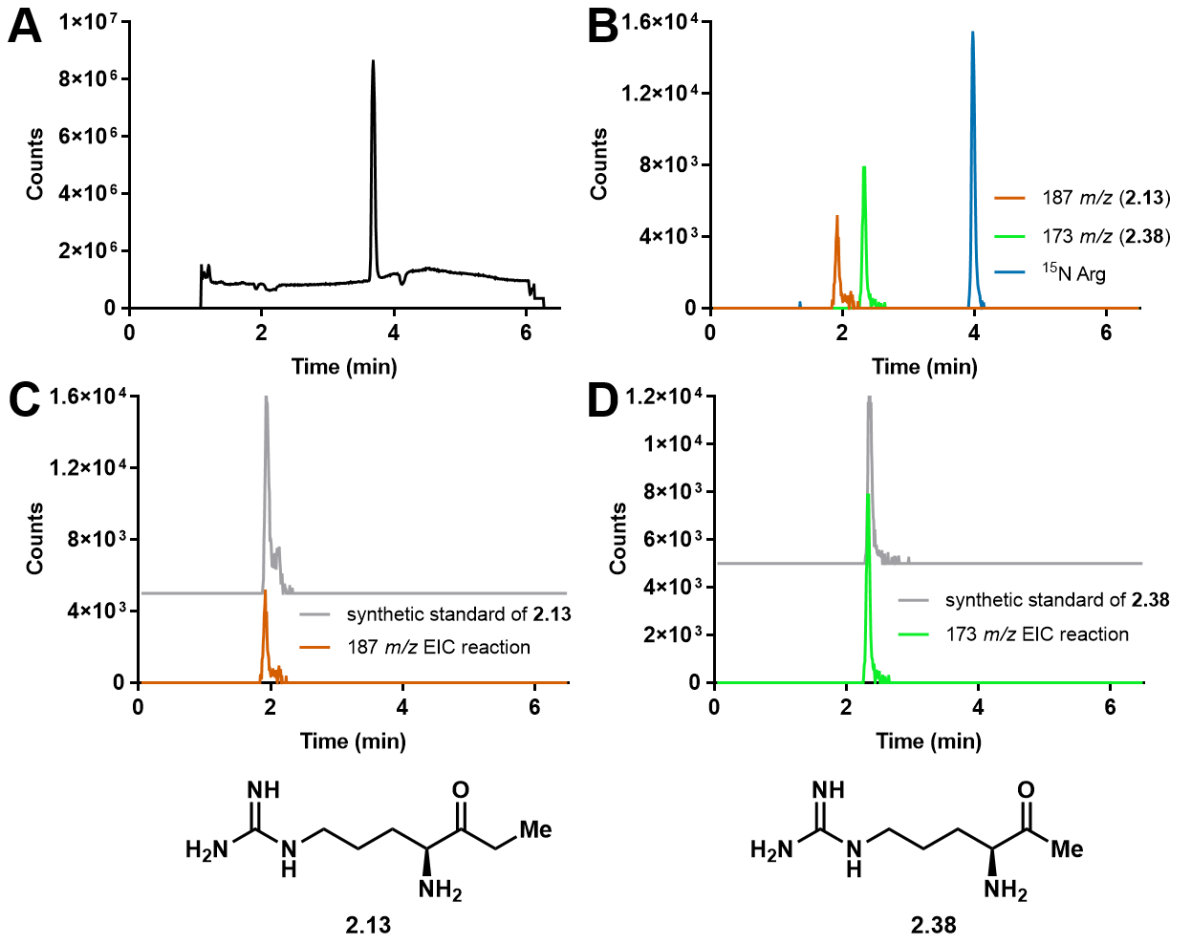


Figure 2.S9. Representative LC-MS analysis of full SxtA module reactions.
(A) Total ion chromatogram; (B) combined extracted ion chromatograms of products ethyl ketone (**2.13**, red), methyl ketone (**2.38**, green), and the ^{15}N Arg internal standard (blue). (C) Extracted ion chromatograms for 187 *m/z* of the synthetic standard of **2.13** (gray) and the reaction (red). (D) Extracted ion chromatograms for 173 *m/z* of the synthetic standard of **2.38** (gray) and the reaction (green).

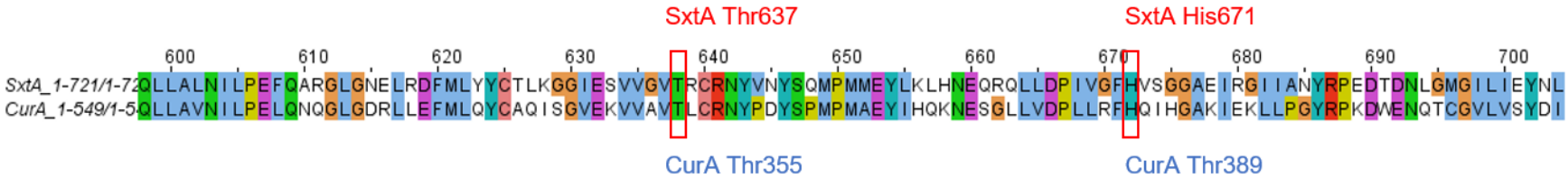
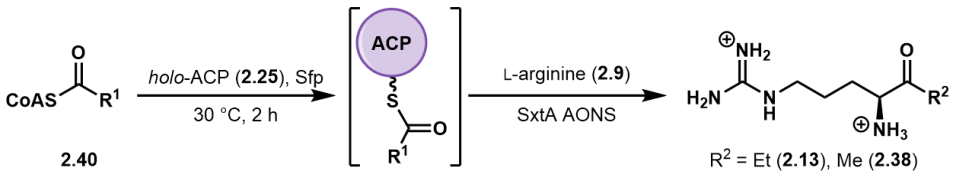
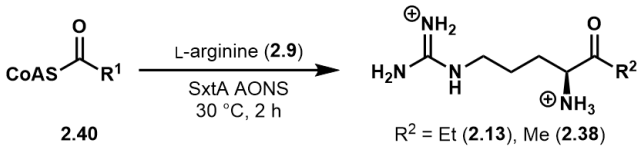


Figure 2.S10. Clustal Omega alignment of the SxtA and CurA GNAT domains.
The catalytic residues of SxtA GNAT are boxed in red.



AONS-mediated condensation of arginine and acyl-ACPs (Table 2.3): 500 μM acyl-CoA, 2 mM Arg, 50 μM *apo*-ACP, 10 μM Sfp, 8 μM ACP(KO)-AONS in 50 mM HEPES pH 7.4, 150 mM NaCl, 1 mM $MgCl_2$, 0.5 mM $MnCl_2$ (total volume 25 μL) were combined and incubated at 30 °C for 2 h. Reactions were quenched by diluting 5x with acetonitrile. Precipitate was pelleted at 12,000 x *g* for 20 min. The supernatant was diluted 20x with the LC-MS analysis mixture for a final ratio of 5% quenched reaction, 2.5% of aqueous 10 μg/mL ^{15}N Arg·HCl (Cambridge Isotope Laboratories), 46% acetonitrile, 46.5% of 1% formic acid in water (e.g., 10 μL of quenched reaction, 5 μL of 10 μg/mL ^{15}N Arg·HCl, 92 μL acetonitrile, 93 μL of 1% formic acid in water for a total volume of 200 μL). Samples were analyzed by LC-MS: column = Waters Acquity 1.7 μm UPLC BEH HILIC 2.1 x 100 mm column; method = 85% B at 0.3 mL/min for 2 min, followed by a linear gradient to 75% B at 0.4 mL/min over 3 min and a second linear gradient to 60% B over 0.5 min, 60% B for 1 min and then 6.5 min re-equilibration at 85% B at 0.3 mL/min (total time 15 min). t_R (^{15}N Arg): 4.0 min. Yields were calculated by comparing the areas under the curve of the product extracted ion chromatograms to the ^{15}N Arg internal standard and normalized to the propionyl-CoA reaction (Table 2.3, entry 2). Reactions were run in triplicate.

The ACP(KO)-AONS didomain was chosen over the standalone AONS due to the increased solubility of the didomain. Only the catalytically active AONS is shown in the scheme above.



AONS-mediated condensation of arginine and acyl-CoAs (Table 2.4): 500 μM acyl-CoA, 2 mM Arg, 8 μM ACP(KO)-AONS in 50 mM HEPES pH 7.4, 150 mM NaCl, 1 mM $MgCl_2$, 0.5 mM $MnCl_2$ (total volume 25 μL) were combined and incubated at 30 °C for 2 h. Reactions were quenched by diluting 5x with acetonitrile. Precipitate was pelleted at 12,000 x *g* for 20 min. The supernatant was diluted 20x with the LC-MS analysis mixture for a final ratio of 5% quenched reaction, 2.5% of aqueous 10 μg/mL ^{15}N Arg·HCl (Cambridge Isotope Laboratories), 46% acetonitrile, 46.5% of 1% formic acid in water (e.g., 10 μL of quenched reaction, 5 μL of 10 μg/mL ^{15}N Arg·HCl, 92 μL acetonitrile, 93 μL of 1% formic acid in water for a total volume of 200 μL). Samples were analyzed by LC-MS: column = Waters Acquity 1.7 μm UPLC BEH HILIC 2.1 x 100 mm column; method = 85% B at 0.3 mL/min for 2 min, followed by a linear gradient to 75% B at 0.4 mL/min over 3 min and a second linear gradient to 60% B over 0.5 min, 60% B for 1 min and then 6.5 min re-equilibration at 85% B at 0.3 mL/min (total time 15 min). t_R (^{15}N Arg): 4.0 min. Yields were calculated by comparing the areas under the curve of the product extracted ion chromatograms to the ^{15}N Arg internal standard and normalized to the propionyl-CoA reaction (Table 2.4, entry 2). Reactions were run in triplicate.

2.7 References

(1) Quayle, D. B. *Paralytic Shellfish Poisoning in British Columbia*; Stevenson, J. C., Ed.; Ottowa, 1969.

(2) Llewellyn, L. E. Saxitoxin, a Toxic Marine Natural Product That Targets a Multitude of Receptors. *Nat. Prod. Rep.* **2006**, *23*, 200–222.

(3) Shinohara, R.; Akimoto, T.; Iwamoto, O.; Hirokawa, T.; Yotsu-Yamashita, M.; Yamaoka, K.; Nagasawa, K. Synthesis of Skeletal Analogues of Saxitoxin Derivatives and Evaluation of Their Inhibitory Activity on Sodium Ion Channels Nav1.4 and Nav1.5. *Chem. Eur. J.* **2011**, *17*, 12144–12152.

(4) Müller, H. The Chemistry and Toxicity of Mussel Poison. *J. Pharmacol. Exp. Ther.* **1935**, *67*, 67–89.

(5) Sommer, H.; Whedon, W. F.; Kofoid, C. A.; Stohler, R. Relation of Paralytic Shell-Fish Poison to Certain Plankton Organisms of the Genus Gonyaulax. *Arch. Pathol.* **1937**, *24*, 537–559.

(6) Sommer, H.; Meyer, K. F. Paralytic Shell-Fish Poisoning. *Arch. Pathol.* **1937**, *24*, 560–598.

(7) Thottumkara, A. P.; Parsons, W. H.; Du Bois, J. Saxitoxin. *Angew. Chem. Int. Ed.* **2014**, *53*, 5760–5784.

(8) Schantz, E. J.; Mold, J. D.; Stanger, D. W.; Shavel, J.; Riel, F. J.; Bowden, J. P.; Lynch, J. M.; Savage Wyler, R.; Riegel, B.; Sommer, H. Paralytic Shellfish Poison. VI. A Procedure for the Isolation and Purification of the Poison from Toxic Clam and Mussel Tissues. *J. Am. Chem. Soc.* **1957**, *79*, 5230–5235.

(9) Stanger, D. W.; Maurer, J. E.; Lynch, J. M.; Schantz, E. J.; Riegel, B. Paralytic Shellfish Poison. VII . Evidence for the Purity of the Poison Isolated from Toxic Clams and Mussels. *J. Am. Chem. Soc.* **1957**, *1015*, 5235–5238.

(10) Schantz, E. J.; Ghazarossian, V. E.; Schnoes, H. K.; Strong, F. M.; Springer, J. P.; Pezzanite, J. O.; Clardy, J. Structure of Saxitoxin. *J. Am. Chem. Soc.* **1975**, *25*, 1238–1239.

(11) Bordner, J.; Thiessen, W. E.; Bates, H. A.; Rapoport, H. Structure of a Crystalline Derivative of Saxitoxin. Structure of Saxitoxin. *J. Am. Chem. Soc.* **1975**, *97*, 6008–6012.

(12) Tanino, H.; Nakata, T.; Kaneko, Y.; Kishi, Y. A Stereospecific Total Synthesis of DL-Saxitoxin. *J. Am. Chem. Soc.* **1977**, *99*, 2818–2819.

(13) Eijkelkamp, N.; Linley, J. E.; Baker, M. D.; Minett, M. S.; Cregg, R.; Werdehausen, R.; Wood, J. N. Neurological Perspectives on Voltage-Gated Sodium Ion Channels. *Brain* **2012**, *135*, 2585–2612.

(14) Mulcahy, J. V; Pajouhesh, H.; Beckley, J. T.; Delwig, A.; Bois, J. Du; Hunter, J. C. Challenges and Opportunities for Therapeutics Targeting the Voltage-Gated Sodium Channel Isoform Na V 1 . 7. *J. Med. Chem.* **2019**, *62*, 8695–8710.

(15) Lattes, K.; Venegas, P.; Lagos, N.; Lagos, M.; Pedraza, L.; Rodriguez-Navarro, A. J.; García, C. Local Infiltration of Gonyautoxin Is Safe and Effective in Treatment of Chronic Tension-Type Headache Local Infiltration of Gonyautoxin Is Safe and Effective in Treatment of Chronic Tension-Type Headache. *Neurol. Res.* **2013**, *31*, 228–233.

(16) Hinzpeter, J.; Palet, M.; Wulf, R.; Barahona, M.; Sepúlveda, J. M.; Guerra, M.; Bustamante, T.; Del, M.; Tapia, E.; Lagos, N. Gonyautoxins: First Evidence in Pain Management in Total Knee Arthroplasty. *Toxicon* **2016**, *119*, 180–185.

(17) Kellmann, R.; Neilan, B. A. Biochemical Characterization of Paralytic Shellfish Toxin Biosynthesis in Vitro. *J. Phycol.* **2007**, *43*, 497–508.

(18) Kellmann, R.; Mihali, T. K.; Young, J. J.; Pickford, R.; Pomati, F.; Neilan, B. A.

Biosynthetic Intermediate Analysis and Functional Homology Reveal a Saxitoxin Gene Cluster in Cyanobacteria. *Appl. Environ. Microbiol.* **2008**, *74*, 4044–4053.
(19) Shimizu, Y. Microalgal Metabolites. *Chem. Rev.* **1993**, *93*, 1685–1698.
(20) Mihali, T. K.; Kellmann, R.; Neilan, B. A. Characterisation of the Paralytic Shellfish Toxin Biosynthesis Gene Clusters in Anabaena Circinalis AWQC131C and Aphanizomenon Sp. NH-5. *BMC Biochem.* **2009**, *10*, 8.
(21) Mihali, T. K.; Carmichael, W. W.; Neilan, B. A. A Putative Gene Cluster from a Lyngbya Wollei Bloom That Encodes Paralytic Shellfish Toxin Biosynthesis. *PLoS ONE* **2011**, *6*, e14657.
(22) Tsuchiya, S.; Cho, Y.; Konoki, K.; Nagasawa, K.; Oshima, Y.; Yotsu-Yamashita, M. Biosynthetic Route towards Saxitoxin and Shunt Pathway. *Sci. Rep.* **2016**, *6*, 20340.
(23) Tsuchiya, S.; Cho, Y.; Konoki, K.; Nagasawa, K.; Oshima, Y.; Yotsu-Yamashita, M. Synthesis and Identification of Proposed Biosynthetic Intermediates of Saxitoxin in the Cyanobacterium Anabaena Circinalis (TA04) and the Dinoflagellate Alexandrium Tamarense (Axat-2). *Org. Biomol. Chem.* **2014**, *12*, 3016–3020.
(24) Vetting, M. W.; de Carvalho, L. P. S.; Yu, M.; Hegde, S. S.; Magnet, S.; Roderick, S. L.; Blanchard, J. S. Structure and Functions of the GNAT Superfamily of Acetyltransferases. *Arch. Biochem. Biophys.* **2005**, *433*, 212–226.
(25) Gu, L.; Geders, T. W.; Wang, B.; Gerwick, W. H.; Hakansson, K.; Smith, J. L.; Sherman, D. H. GNAT-Like Strategy for Polyketide Chain Initiation. *Science* **2007**, *318*, 970–975.
(26) Skiba, M. A.; Sikkema, A. P.; Moss, N. A.; Tran, C. L.; Sturgis, R. M.; Gerwick, L.; Gerwick, W. H.; Sherman, D. H.; Smith, J. L. A Mononuclear Iron-Dependent Methyltransferase Catalyzes Initial Steps in Assembly of the Apratoxin A Polyketide Starter Unit. *ACS Chem. Biol.* **2017**, *12*, 3039–3048.
(27) Skiba, M. A. Structural and Biochemical Investigation of Methylation and Elucidation of T-Butyl Formation in Polyketide Biosynthesis, University of Michigan, 2018.
(28) Broderick, J. B.; Duffus, B. R.; Duschene, K. S.; Shepard, E. M. Radical S-Adenosylmethionine Enzymes. *Chem. Rev.* **2014**, *114*, 4229–4317.
(29) Simunovic, V.; Zapp, J.; Rachid, S.; Krug, D.; Meiser, P.; Müller, R. Myxovirescin A Biosynthesis Is Directed by Hybrid Polyketide Synthases/Nonribosomal Peptide Synthetase, 3-Hydroxy-3-Methylglutaryl-CoA Synthases, and Trans-Acting Acyltransferases. *ChemBioChem* **2006**, *7*, 1206–1220.
(30) Young, J.; Stevens, D. C.; Carmichael, R.; Tan, J.; Rachid, S.; Boddy, C. N.; Müller, R.; Taylor, R. E. Elucidation of Gephyronic Acid Biosynthetic Pathway Revealed Unexpected SAM-Dependent Methylations. *J. Nat. Prod.* **2013**, *76*, 2269–2276.
(31) Sudek, S.; Lopanik, N. B.; Waggoner, L. E.; Hildebrand, M.; Anderson, C.; Liu, H.; Patel, A.; Sherman, D. H.; Haygood, M. G. Identification of the Putative Bryostatin Polyketide Synthase Gene Cluster from "Candidatus Endobugula Sertula", the Uncultivated Microbial Symbiont of the Marine Bryozoan Bugula Neritina. *J. Nat. Prod.* **2007**, *70*, 67–74.
(32) Grindberg, R. V.; Ishoey, T.; Brinza, D.; Esquenazi, E.; Coates, R. C.; Liu, W.; Gerwick, L.; Dorrestein, P. C.; Pevzner, P.; Lasken, R.; Gerwick, W. H. Single Cell Genome Amplification Accelerates Identification of the Apratoxin Biosynthetic Pathway from a Complex Microbial Assemblage. *PLoS ONE* **2011**, *6*, e18565.
(33) Keatinge-Clay, A. T. The Uncommon Enzymology of Cis-Acyltransferase Assembly Lines. *Chem. Rev.* **2017**, *117*, 5334–5366.
(34) Skiba, M. A.; Sikkema, A. P.; Moss, N. A.; Lowell, A. N.; Su, M.; Sturgis, R. M.; Gerwick,

L.; Gerwick, W. H.; Sherman, D. H.; Smith, J. L. Biosynthesis of T-Butyl in Apratoxin A: Functional Analysis and Architecture of a PKS Loading Module. *ACS Chem. Biol.* **2018**, *13*, 1640–1650.
(35) Callahan, B.; Thattai, M.; Shraiman, B. I. Emergent Gene Order in a Model of Modular Polyketide Synthases. *Proc. Natl. Acad. Sci.* **2009**, *106*, 19410–19415.
(36) Remington, S. J. Mechanisms of Citrate Synthase and Related Enzymes (Triose Phosphate Isomerase and Mandelate Racemase). *Curr. Opin. Struct. Biol.* **1992**, *2*, 730–735.
(37) Poust, S.; Phelan, R. M.; Deng, K.; Katz, L.; Petzold, C. J.; Keasling, J. D. Divergent Mechanistic Routes for the Formation of Gem-Dimethyl Groups in the Biosynthesis of Complex Polyketides. *Angew. Chem. Int. Ed.* **2015**, *54*, 2370–2373.
(38) Alexeev, D.; Alexeeva, M.; Baxter, R. L.; Campopiano, D. J.; Webster, S. P.; Sawyer, L. The Crystal Structure of 8-Amino-7-Oxononanoate Synthase: A Bacterial PLP-Dependent, Acyl-CoA-Condensing Enzyme. *J. Mol. Biol.* **1998**, *284*, 401–419.
(39) Webster, S. P.; Alexeev, D.; Campopiano, D. J.; Watt, R. M.; Alexeeva, M.; Sawyer, L.; Baxter, R. L. Mechanism of 8-Amino-7-Oxononanoate Synthase: Spectroscopic, Kinetic, and Crystallographic Studies. *Biochemistry* **2000**, *39*, 516–528.
(40) Pfeifer, B. A.; Admiraal, S. J.; Gramajo, H.; Cane, D. E.; Khosla, C. Biosynthesis of Complex Polyketides in a Metabolically Engineered Stain of E. Coli. *Science* **2001**, *291*, 1790–1792.
(41) Roulet, J.; Taton, A.; Golden, J. W.; Arabolaza, A.; Burkart, M. D.; Gramajo, H. Development of a Cyanobacterial Heterologous Polyketide Production Platform. *Metab. Eng.* **2018**, *49*, 94–104.
(42) Liu, X.; Sheng, J.; Curtiss III, R. Fatty Acid Production in Genetically Modified Cyanobacteria. *Proc. Natl. Acad. Sci.* **2011**, *108*, 6899–6904.
(43) Eschenfeldt, W. H.; Lucy, S.; Millard, C. S.; Joachimiak, A.; Mark, I. D. A Family of LIC Vectors for High-Throughput Cloning and Purification of Proteins. In *Methods in Molecular Biology*; Doyle, S. A., Ed.; Humana Press: Totowa, NJ, 2009; Vol. 498, pp 105–115.
(44) Dorrestein, P. C.; Bumpus, S. B.; Calderone, C. T.; Garneau-Tsodikova, S.; Aron, Z. D.; Straight, P. D.; Kolter, R.; Walsh, C. T.; Kelleher, N. L. Facile Detection of Acyl and Peptidyl Intermediates on Thiotemplate Carrier Domains via Phosphopantetheinyl Elimination Reactions during Tandem Mass Spectrometry. *Biochemistry* **2006**, *45*, 12756–12766.
(45) Hitchman, T. S.; Crosby, J.; Byrom, K. J.; Cox, R. J.; Simpson, T. J. Catalytic Self-Acylation of Type II Polyketide Synthase Acyl Carrier Proteins. *Chem. Biol.* **1998**, *5*, 35–47.
(46) Arthur, C. J.; Szafranska, A. E.; Long, J.; Mills, J.; Cox, R. J.; Findlow, I. S. C.; Simpson, T. J.; Crump, M. P.; Crosby, J. The Malonyl Transferase Activity of Type II Polyketide Synthase Acyl Carrier Proteins. *Chem. Biol.* **2006**, *13*, 587–596.
(47) Florova, G.; Kazanina, G.; Reynolds, K. A. Enzymes Involved in Fatty Acid and Polyketide Biosynthesis in Streptomyces Glaucescens: Role of FabH and FabD and Their Acyl Carrier Protein Specificity. *Biochemistry* **2002**, *41*, 10462–10471.
(48) Dreier, J.; Shah, A. N.; Khosla, C. Kinetic Analysis of the Actinorhodin Aromatic Polyketide Synthase. *J. Biol. Chem.* **1999**, *274*, 25108–25112.
(49) Chun, S. W.; Hinze, M. E.; Skiba, M. A.; Narayan, A. R. H. Chemistry of a Unique Polyketide-like Synthase. *J. Am. Chem. Soc.* **2018**, *140*, 2430–2433.

(50) Skiba, M. A.; Tran, C. L.; Dan, Q.; Sikkema, A. P.; Klaver, Z.; Gerwick, W. H.; Sherman, D. H.; Smith, J. L. Repurposing the GNAT Fold in the Initiation of Polyketide Biosynthesis. *Structure* **2020**, *28*, 63–74.

(51) Froese, D. S.; Forouhar, F.; Tran, T. H.; Vollmar, M.; Kim, Y. S.; Lew, S.; Neely, H.; Seetharaman, J.; Shen, Y.; Xiao, R.; Acton, T. B.; Everett, J. K.; Cannone, G.; Puranik, S.; Savitsky, P.; Krojer, T.; Pilka, E. S.; Kiyani, W.; Lee, W. H.; Marsden, B. D.; Delft, F. Von; Allerston, C. K.; Spagnolo, L.; Gileadi, O.; Montelione, G. T.; Oppermann, U.; Yue, W. W.; Tong, L. Crystal Structures of Malonyl-Coenzyme A Decarboxylase Provide Insights into Its Catalytic Mechanism and Disease-Causing Mutations. *Struct. Des.* **2013**, *21*, 1182–1192.

(52) Aparicio, D.; Pérez-Luque, R.; Carpena, X.; Díaz, M.; Ferrer, J. C.; Loewen, P. C.; Fita, I. Structural Asymmetry and Disulfide Bridges among Subunits Modulate the Activity of Human Malonyl-CoA Decarboxylase. *J. Biol. Chem.* **2013**, *288*, 11907–11919.

(53) Skiba, M. A.; Sikkema, A. P.; Fiers, W. D.; Gerwick, W. H.; Sherman, D. H.; Aldrich, C. C.; Smith, J. L. Domain Organization and Active Site Architecture of a Polyketide Synthase C-Methyltransferase. *ACS Chem. Biol.* **2016**, *11*, 3319–3327.

(54) Quadri, L. E. N.; Weinreb, P. H.; Lei, M.; Nakano, M. M.; Zuber, P.; Walsh, C. T. Characterization of Sfp, a Bacillus Subtilis Phosphopantetheinyl Transferase for Peptidyl Carrier Protein Domains in Peptide Synthetases. *Biochemistry* **1998**, *37*, 1585–1595.

(55) Heckman, K. L.; Pease, L. R. Gene Splicing and Mutagenesis by PCR-Driven Overlap Extension. *Nat. Protoc.* **2007**, *2*, 924–932.

(56) Sánchez, C.; Du, L.; Edwards, D. J.; Toney, M. D.; Shen, B. Cloning and Characterization of a Phosphopantetheinyl Transferase from Streptomyces Verticillus ATCC15003, the Producer of the Hybrid Peptide–Polyketide Antitumor Drug Bleomycin. *Chem. Biol.* **2001**, *8*, 725–738.

(57) Gräslund, S.; Nordlund, P.; Weigelt, J.; Bray, J.; Gileadi, O.; Hallberg, B. M.; Knapp, S.; Oppermann, U.; Arrowsmith, C.; Hui, R.; Ming, J.; Dhe-Paganon, S.; Park, H.; Savchenko, A.; Yee, A.; Edwards, A.; Vincentelli, R.; Cambillau, C.; Kim, R.; Kim, S.-H.; Rao, Z.; Shi, Y.; Terwilliger, T. C.; Kim, C.-Y.; Hung, L.-W.; Waldo, G. S.; Peleg, Y.; Albeck, S.; Unger, T.; Dym, O.; Prilusky, J.; Sussman, J. L.; Stevens, R. C.; Lesley, S. A.; Wilson, I. A.; Joachimiak, A.; Collart, F.; Dementieva, I.; Donnelly, M. I.; Eschenfeldt, W. H.; Kim, Y.; Stols, L.; Wu, R.; Zhou, M.; Burley, S. K.; Emtage, J. S.; Sauder, J. M.; Thompson, D.; Bain, K.; Luz, J.; Gheyi, T.; Zhang, F.; Atwell, S.; Almo, S. C.; Bonanno, J. B.; Fiser, A.; Swaminathan, S.; Studier, F. W.; Chance, M. R.; Sali, A.; Acton, T. B.; Xiao, R.; Zhao, L.; Ma, L. C.; Hunt, J. F.; Tong, L.; Cunningham, K.; Inouye, M.; Anderson, S.; Janjua, H.; Shastry, R.; Ho, C. K.; Wang, D.; Wang, H.; Jiang, M.; Montelione, G. T.; Stuart, D. I.; Owens, R. J.; Daenke, S.; Schütz, A.; Heinemann, U.; Yokoyama, S.; Büssow, K.; Gunsalus, K. C. Protein Production and Purification. *Nat. Methods* **2008**, *5*, 135–146.

Chapter 3: Chemoenzymatic Synthesis of α-Amino Ketones with SxtA AONS

Summary

α-amino ketones are an important building block in pharmaceutical products and total synthesis efforts. In contrast to the 4-6 step synthetic routes to stereospecifically prepare α-amino ketones from α-amino acids, the analogous transformations mediated by α-oxoamine synthases (AOS) are a single step, under mild protecting group-free conditions. We recently demonstrated that one AOS enzyme, SxtA AONS (8-amino-7-oxononanoate synthase) affords α-amino ketones from the amino acid L-Arg and acyl-ACP or -CoA thioesters. This Chapter summarizes our work to develop SxtA AONS as a biocatalyst to synthesize α-amino ketones chemoenzymatically on preparative-scale. Reactions were optimized by leveraging the SxtA ACP-AONS didomain as a carrier protein-partner enzyme platform and thiophenol esters as inexpensive acyl donors. However, we also observed that SxtA AONS is only active for fewer than 4 h and approximately 30 turnovers, which is too low for preparative-scale reactions. The final stereoselectivity of the AONS-mediated condensation reaction therefore cannot be determined yet. Additionally, homologs of SxtA ACP-AONS were screened with other amino acids and acyl thioesters to interrogate SxtA AONS' ability

to synthesize a variety of ketones. The enzyme operated on several non-native substrates, but with very low conversion. To evolve variants that operate on the indole-containing tryptophan, we began collaborating with the group of Prof. Robert Kennedy at the University of Michigan on a protein engineering campaign. Two potential hits in a first-generation random mutagenesis library were found, and efforts to further engineer this protein are ongoing.

3.1 Introduction

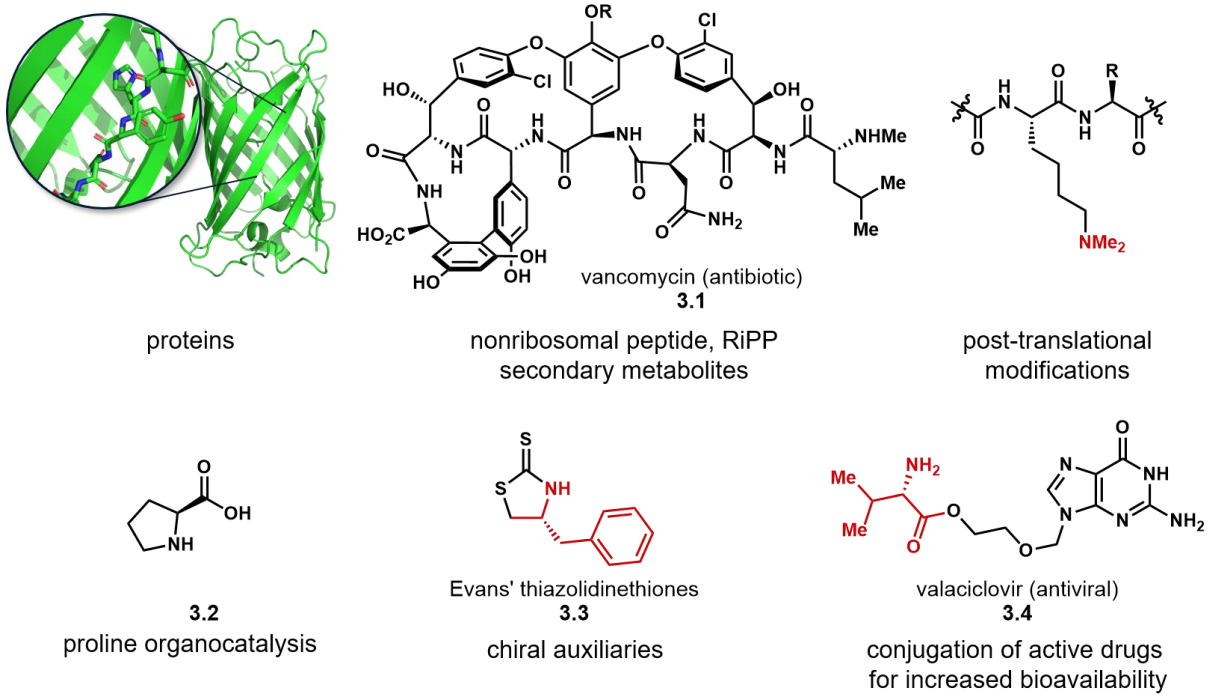


Figure 3.1. Uses of α-amino acids in biology and chemistry.

Amino acids serve many important roles in biology, where they are primarily the building blocks of proteins, but also incorporated into nonribosomal peptide and ribosomally synthesized and post-translationally modified peptide (RiPP) natural products (Figure 3.1).[1,2] In synthetic chemistry, amino acids are employed as catalysts and precursors to chiral auxiliaries (**3.3**) and pharmaceutical compounds.[3–6] The twenty proteinogenic amino acids are an economical class of chiral pool molecules, and a diverse array of non-proteinogenic amino acids are also commercially available.[7,8]

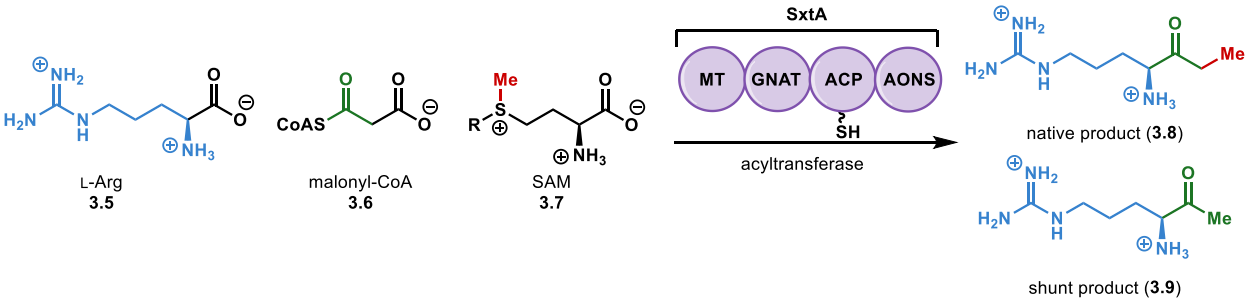


Figure 3.2. The polyketide-like synthase SxtA natively mediates assembly of ketone 3.8.

Free amino acids and protein residues are often modified to adjust the structure and function of the final products;[1,9,10] there are thousands of enzymes that mediate these transformations endogenously, but synthetic methods to derivatize amino acid functions are fewer and generally more challenging.[11] The amino and carboxylic acid motifs usually require protection, as do reactive side chains, adding more steps to a synthetic sequence.[12] Therefore, the development of synthetic methods that are selective and protecting group-free would be valuable.

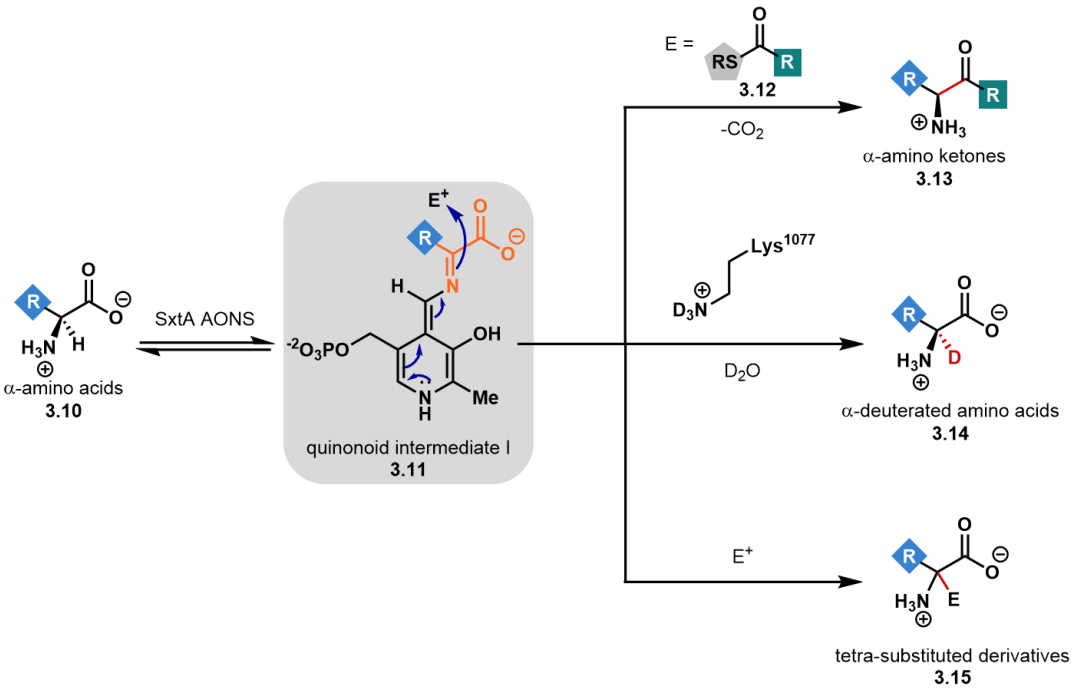


Figure 3.3. AONS-catalyzed diversification of α-amino acids stemming from a nucleophilic intermediate.

Based on our previous experiments described in Chapter 2 demonstrating that SxtA AONS converted the amino acid L-Arg to α-amino ketones in a single step, we envisioned developing the domain as a general biocatalyst to modify amino acids. The key reactive intermediate in the mechanism of SxtA AONS is the nucleophilic quinonoid I (see **3.11**, Figure 3.3).[13,14] In the native reaction, quinonoid **3.11** attacks an electrophilic acyl-ACP thioester, eventually affording three products: ketone **3.13**, carbon dioxide, and *holo*-ACP. Replacing thioesters with other

electrophiles, would lead to other types of derivatized products, such as α-deuterated amino acids (**3.14**, discussed in Chapter 4), or tetra-substituted molecules (**3.15**).

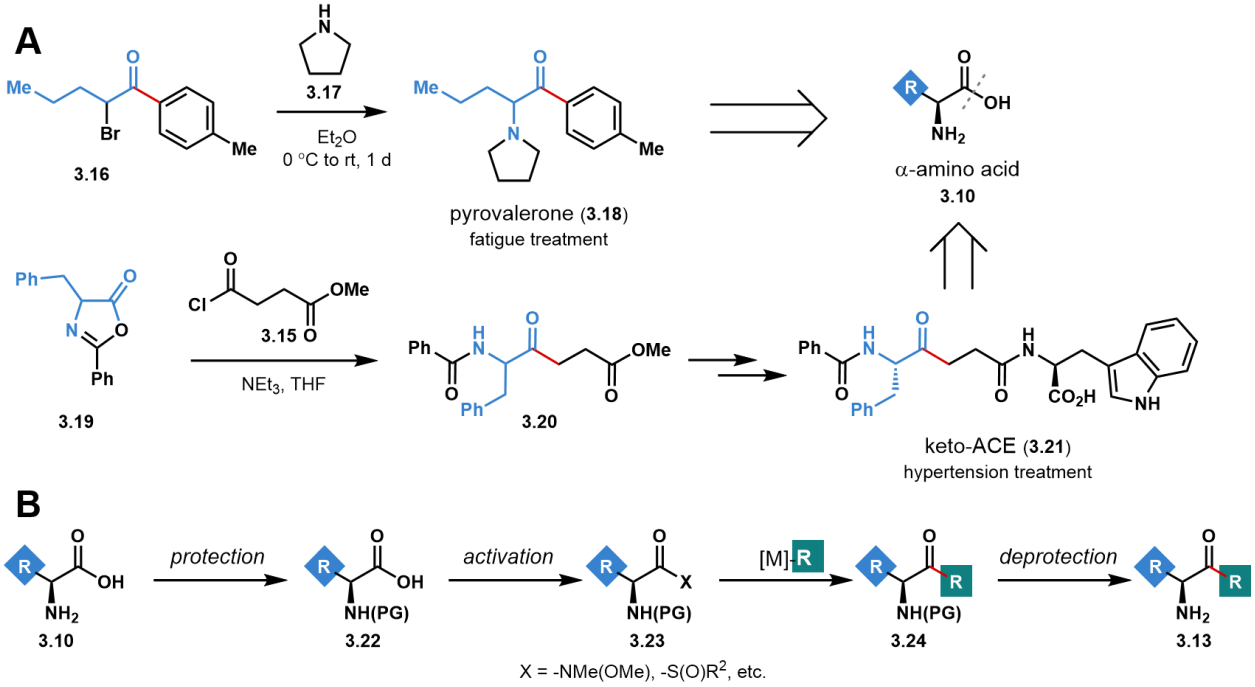


Figure 3.4. Synthetic methods toward α-amino ketones.
(A) α-amino ketone-forming steps in pharmaceutical products are not stereoselective and proceed but could potentially be synthesized from an amino acid precursor.[15–17] (B) Stereoselective transformation of an α-amino acid to a ketone requires a multi-step synthetic route using traditional chemical methods.[18–23]

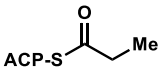
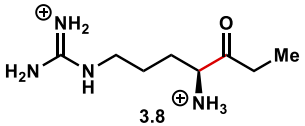
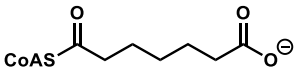
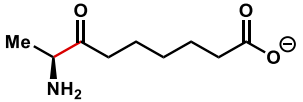
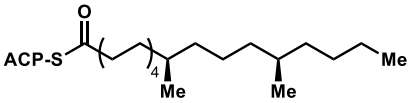
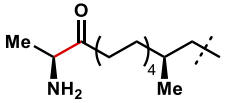
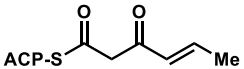
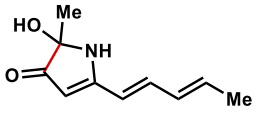
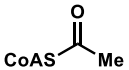
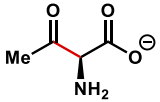
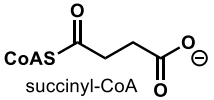
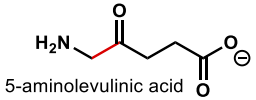
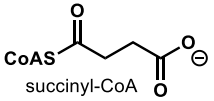
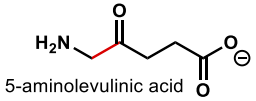
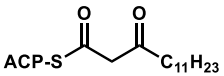
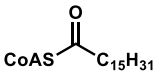
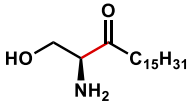
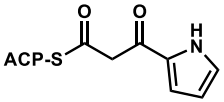
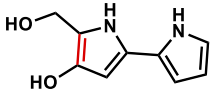
We chose to develop the native reaction type of SxtA AONS, a decarboxylative Claisen-like condensation from an amino acid to a ketone, first. α-amino ketones are valuable precursors in the total syntheses of natural products,[18,24] and are also found in some pharmaceutical products which can be synthesized through various methods (Figure 3.4A). Pyrovalerone (**3.18**) was initially prepared from α-bromo ketone **3.16**, without control over the enantioselectivity, despite only the 2*S*-isomer being bioactive.[15,16] The early syntheses of Keto-ACE (**3.21**) also used the non-selective Dakin-West reaction to synthesize the α-amino ketone segment.[17] The more intuitive option is synthesis from the corresponding amino acids, forming the bonds highlighted in red in Figure 3.4A. These have their own set of drawbacks, including protection of the amino group, and correct

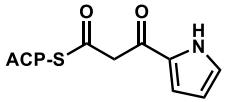
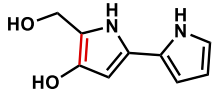
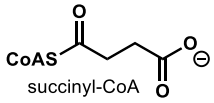
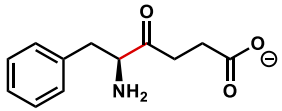
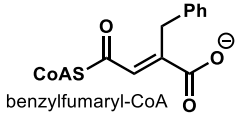
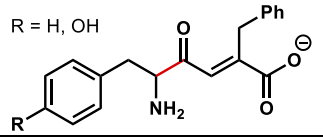
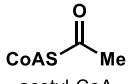
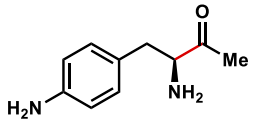
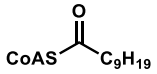
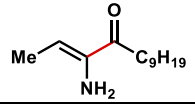
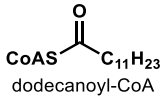
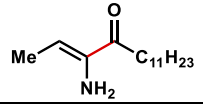
activation of the carbonyl such that any carbon nucleophiles only add once (Figure 3.4B).[18–23] These considerations all add to the step counts, leading to longer inefficient synthetic sequences. As an example, the synthetic standards of SxtA's native full-length product and one shunt product, ketones **3.8** and **3.9**, were prepared in six steps from L-ornithine, compared to the one-step SxtA AONS transformation.[14,25]

The α-oxoamine synthase (AOS) enzyme class to which SxtA AONS belongs is a potential biocatalytic solution to this synthetic challenge, mediating ketone formation that is likely stereospecific. The protein class has been studied for decades for links to human diseases or as potential drug targets.[26–28] There is a common accepted mechanism (see Figure 2.7), but many finer details about AOSs have not been elucidated yet.[29] The substrate scopes of the sixteen known AOS members (Table 3.1) have not been exhaustively probed. Chapter 2 describes preliminary reactions with SxtA AONS, demonstrating that it can employ multiple thiol (ACP, CoA) and acyl group (propionyl, acetyl) sources to enzymatically synthesize **3.8** and **3.9** *in vitro*, suggesting that the domain may also catalyze ketone formation from other non-native combinations of amino acids and thioesters.

In this Chapter, we discuss our efforts to optimize experiments with SxtA AONS for preparative-scale experiments and verify the stereoselectivity of the overall ketone-forming reaction (section 3.2). Additionally, we examined the promiscuity of wild-type SxtA AONS and began a directed evolution campaign to discover a variant protein that operates more on the non-native substrate L-tryptophan (section 3.3).

Table 3.1. Summary of known AOS enzymes

Protein name(s)	Biosynthetic pathway	Amino acid substrate	Thioester	Product	Notes
SxtA AONS[14] (domain 4)	Saxitoxin	L-Arg	 propionyl-ACP	 3.8	
BioF or 8-amino-7-oxononanoate synthase[30]	Biotin	L-Ala	 pimeloyl-CoA		Some species use ACP thioesters[31]
FUM8[32,33]	Fumonisin B1	L-Ala			
CuaB[34]	Curvulamine	L-Ala			CuaB also appears to be an oxygenase; released product cyclizes
2-amino-3-ketobutyric acid ligase (KBL)[35]	L-Thr	Gly	 acetyl-CoA		No decarboxylation
HemA or aminolevulinic acid synthase (ALAS)[26,36]	Porphyrins	Gly	 succinyl-CoA	 5-aminolevulinic acid	
ORF34, AsuA[37,38]	C_5N heterocycles	Gly	 succinyl-CoA	 5-aminolevulinic acid	Also have downstream cyclization activity
RedL domain 6[39,40]	Prodigiosins	Gly			Released product cyclizes
Serine palmitoyltransferase (SPT)[41]	Sphingosines	L-Ser	 palmitoyl-CoA		
PigH and RedN domain 3[42,43]	Prodigiosins and prodiginines, respectively	L-Ser			Released product cyclizes

TamD domain 2[44]	Tamjamine	L-Ser			Released product cyclizes
PqrA[45]	Perquinolines	L-Phe			
KtmB[46]	Ketomemicins and arphamenines	L-Phe and L-Tyr			
PapD[47]	Pyrazines	L-4-amino-Phe			
CqsA[48,49]	CAI-1	SAM			Eliminates (*S*)-methyl-5'-thioadenosine first
LqsA[50]	LAI-1	SAM?			
JqsA[51]	JAI-1	Unknown	Unknown	Unknown	JAI-1 final product structure is unknown

3.2 Optimization for Preparative-Scale Reactions of the Condensation Reaction to Arginine-Based Ketones

Like the partner enzymes described in Chapter 1, the major barrier to leveraging those enzymes and AOS proteins on a synthetically useful scale is the cost of the activated building blocks. Both routes toward the native SxtA product ketone **3.8** we previously investigated in Chapter 2 originate from a CoA source (Figure 3.5). These thioesters are commercially available but extremely expensive compared to amino acids, and not atom-economical due to the high molecular weight of the CoASH byproduct. We first sought to identify the propionyl thioester with the best balance between cost and activity with SxtA AONS.

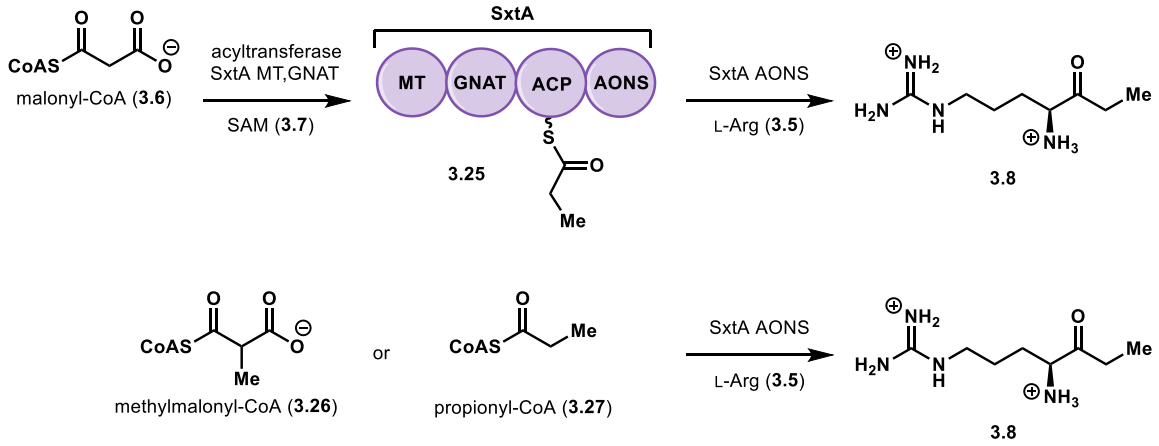


Figure 3.5. Biosynthesis of ketone 3.8 from CoA thioesters.

Reaction Optimization

Eight propionyl thioesters, prepared from thiols of varying cost and size were tested. *N*-acetylcysteamine (SNAC), which resembles the Ppant cofactor of *holo*-ACP, was synthesized by a coupling reaction to propionic acid (Figure 3.6). The thiophenol ester (**3.37**) was easily synthesized in one step from thiophenol and propionic anhydride (Figure 3.6B).[52] The product was then used to prepare the pantetheine and ACP thioesters (**3.38** and **3.39**) by aqueous

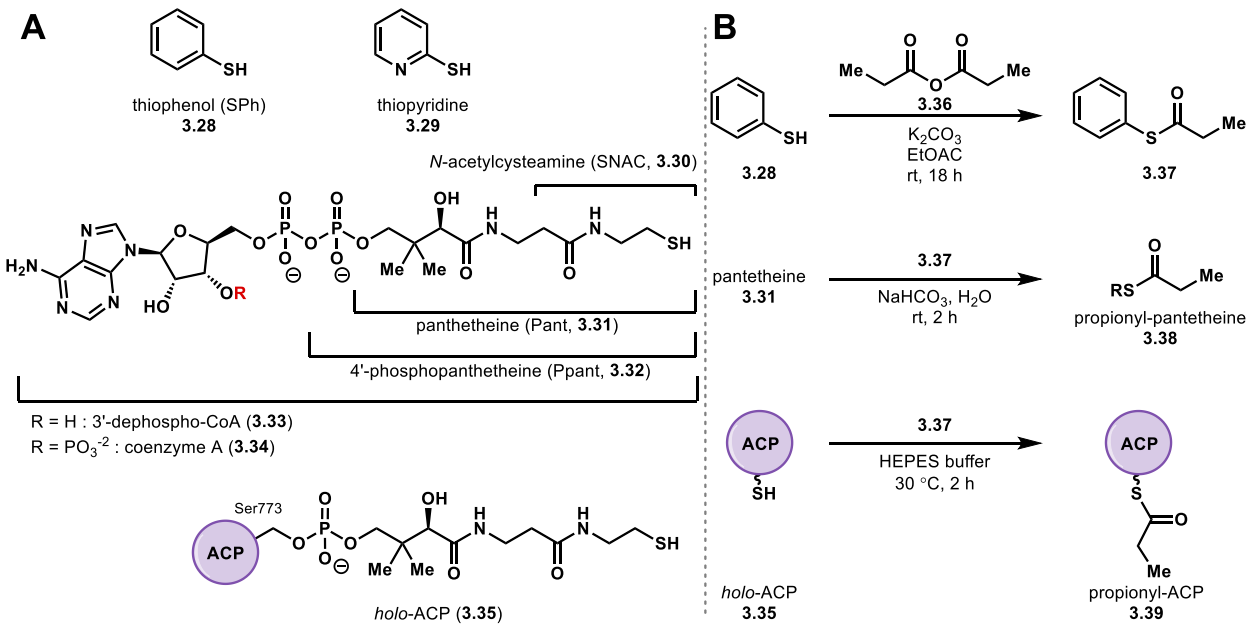


Figure 3.6. Thiol activating groups tested with SxtA AONS.
(A) Comparison of thiols to the native activator, *holo*-ACP. (B) Synthesis of selected propionyl thioesters.

transthioesterification,[53] with separation of the excess small molecules on a PD-10 desalting and size-exclusion column for the protein (about 80% loading, as measured by intact protein MS). Next, thioesters **3.31**-**3.34** were enzymatically generated *in situ* by the CoA biosynthetic enzymes CoaADE: phosphopantetheine **3.31** (Ppant) using CoaA only, 3'-desphospho-CoA **3.32** by CoaAD and CoA **3.33** (commercially available but here was formed by CoaADE).[54]

After running the CoA biosynthesis reactions for 3 hours, L-Arg and SxtA AONS were incubated with the thioesters for an additional 4 hours, and the relative yields of **3.8** were compared (Figure 3.7). At this time, we were using SxtA AONS constructs from *C. raciborskii*, and found that the simplest thiols (thiophenol, thiopyridine, SNAC) were not directly compatible with the standalone domain. In contrast, the pantetheine thioester **3.38** did lead to the desired product formation, indicating that pantetheine was the minimum thiol activator. The CoA thioester positive control afforded **3.8**, as did the CoA biosynthetic intermediates Ppant and 3'-dephospho-CoA. The native thiol activator *holo*-ACP was the highest converting, as expected.

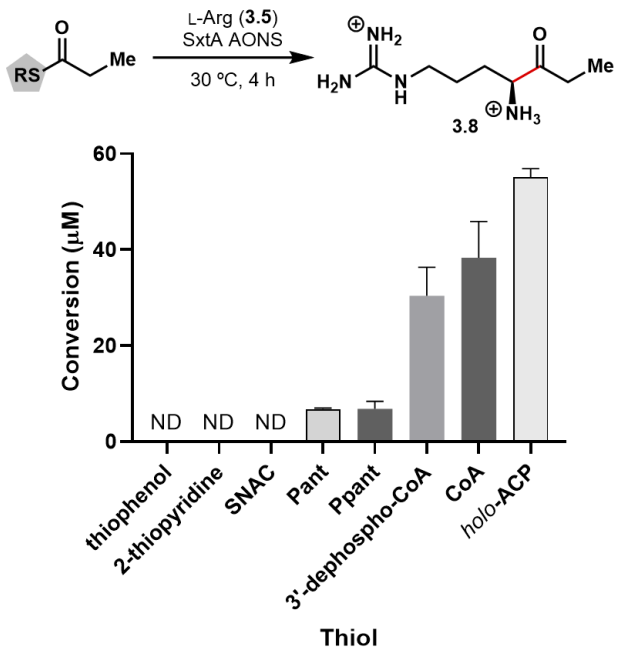


Figure 3.7. Activity test of SxtA AONS with various activating thiols to synthesize ketone 3.8.
Conditions: 250 μM propionyl-pantetheine (**3.38**) or propionyl thioester, 2 mM Arg, 8 mM ATP, 0.2 mg/mL CoaADE.

Although the most cost-effective thiol activators thiophenol, thiopyridine and SNAC were not substrates, employing stoichiometric CoA or *holo*-ACP is still economically unfeasible. However, thiophenol and SNAC thioesters are known to transesterify *holo*-ACPs as well;[55,56] therefore, these could serve as the indirect but inexpensive acyl donors, generating the active acyl-ACP or -CoA *in situ* and allowing the most expensive component, *holo*-ACP or CoASH, to be used in catalytic amounts.

The stoichiometry between ACP and AONS was tested next (Figure 3.8). The relative protein concentration is more easily controlled with standalone *trans*-acting domains, but the *cis*-acting ACP-AONS didomain more closely resembles the native context of SxtA. Of the two possible acyl donors, propionyl-SPh **3.37** afforded a higher relative conversion of ketone **3.8** than the SNAC thioester (**3.40**). Increasing the concentration of *holo*-ACP from 1 equivalent to 6.3 equivalents of

AONS also increased conversion, but the didomain still had the highest relative yield. Adding the thiol scavenger 2-vinylpyridine (2-VP)[57] to react with the thiol byproducts significantly suppressed conversion.

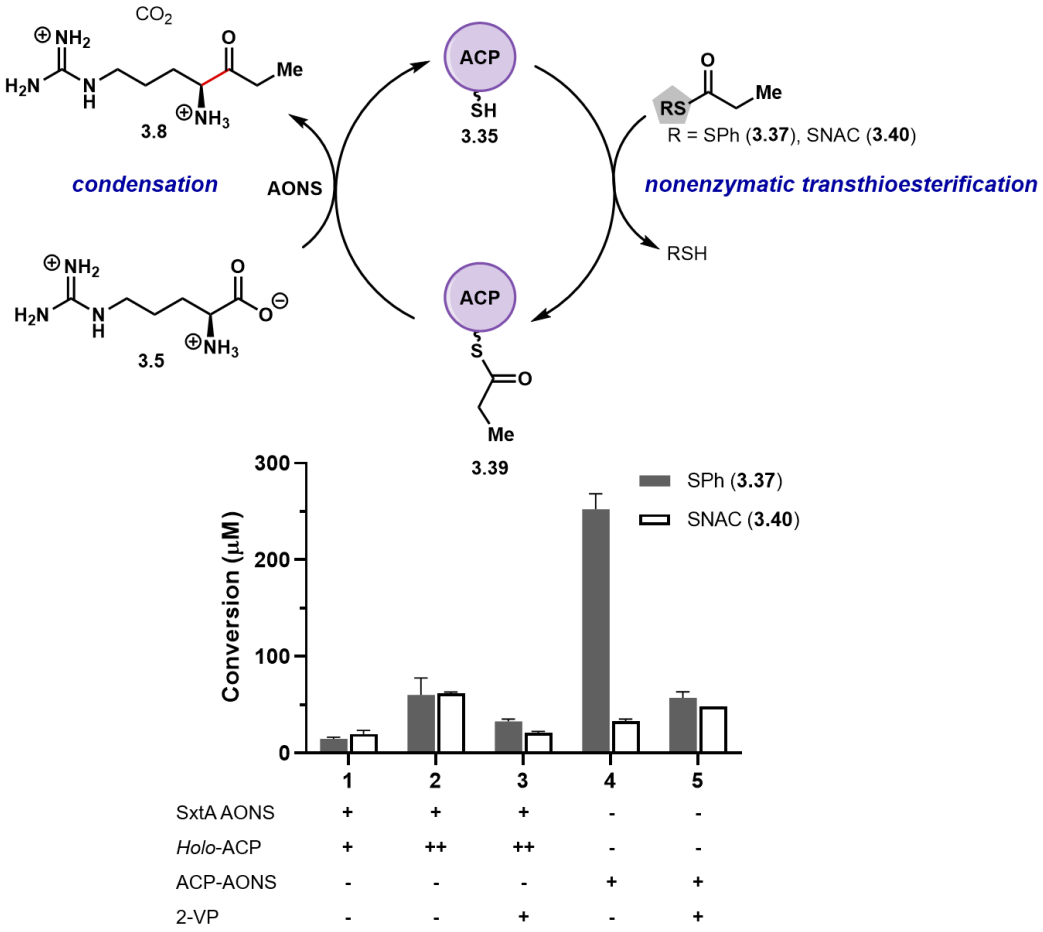
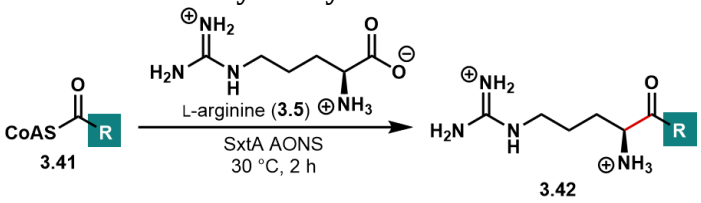


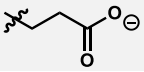
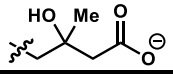
Figure 3.8. Optimization of ACP and AONS stoichiometry with small molecule acyl donors generating propionyl-ACP *in situ*.
+: the catalyst or reagent was present; -: the catalyst or reagent was absent. For *holo*-ACP, + indicates 1 equivalent of protein, and 6.3 equivalents were added in reactions labeled ++. Conditions: 2 mM propionyl thioesters, 4 mM 2-vinylpyridine, 2 mM Arg, 8 μM *C. raciborskii* ACP(KO)-AONS or ACP-AONS, 8 or 50 μM *holo*-ACP.

This semi-optimized reaction format was used to test the acyl group substrate scope of AONS. A panel of ten commercially available acyl-CoA thioesters was screened initially, eight of which formed the corresponding arginine ketones (Table 3.2). Propionyl- and acetyl-CoA gave the ethyl **3.8** and methyl ketones **3.9**, respectively, as described previously in Chapter 2 (entries 1 and 2).

Other straight-chain alkyl groups were accepted (entries 3-5), but the relative yields of the branched chain substrates isobutyryl and isovaleryl (entries 6 and 7) were low. The case of isobutyryl-CoA was particularly illuminating. If the isobutyryl group were formed in our SxtA module experiments (either by dimethylation of malonyl-CoA/-ACP by the SxtA MT domain or monomethylation of the methylmalonyl thioester), it is difficult for the AONS domain to process this substrate. There was slightly more conversion for the larger but more rigid benzoyl-CoA (entry 8). The two polar or negatively charged substrates, succinyl-CoA and 3-hydroxymethylglutaryl- were not productive for ketone formation (entries 9 and 10).

Table 3.2. Activity of acyl-CoA thioesters with SxtA AONS



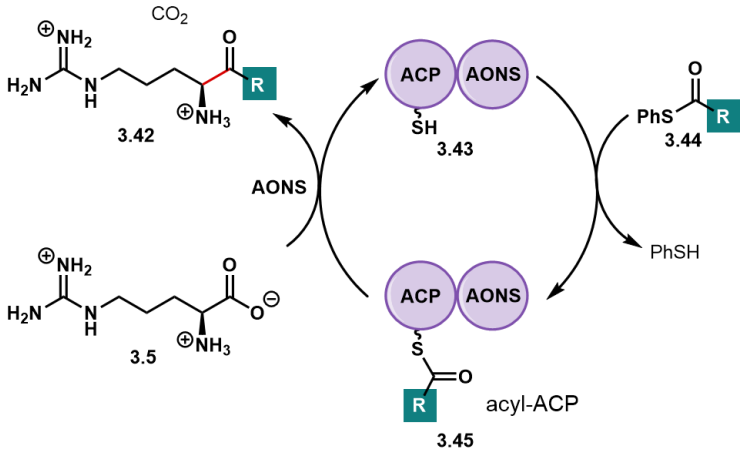
Entry	R =	Normalized yield[a]
1	Me	11 ± 2
2	Et	100 ± 17
3	*n*-Pr	190 ± 10
4	*n*-pentyl	9.2 ± 0.6
5	*n*-heptyl	67 ± 5
6	iPr	2.3 ± 0.5
7	iBu	39 ± 2
8	Ph	2.9 ± 0.2
9		ND
10		ND

ND: not detected. [a]Yields of ketones were calculated by comparing the area under the curve of extracted ion chromatograms to an internal standard, ^{15}N Arg, and normalized to the native product (entry 2).

Employing these CoA thioester substrates is still too expensive for preparative-scale reactions, motivating a screen of so the thiophenol esters of the eight converted acyl groups, as well as

cyclopentyl, cyclohexyl and glycyl substrates. In reactions with ACP-AONS and L-Arg, nine of the eleven substrates were converted to arginine ketones (Table 3.3), generally with a similar reactivity trend as the CoA thioesters. The cyclopropyl group (entry 9) was active with SxtA AONS, but the slightly larger cyclohexyl group (entry 10) and the charged glycine-based aminoacyl group (entry 11) showed no conversion. When the thiophenol esters were incubated separately with *holo*-ACP for 1 h to inspect the loading levels by intact protein MS, the amount of

Table 3.3. Activity of SxtA ACP-AONS with acyl-ACP generated *in situ* from thiophenol esters.



Entry	R =	Normalized yield[a]	Loading of *holo*-ACP
1	Me	16 ± 2	94%
2	Et	100 ± 4	89%
3	*n*-Pr	8.3 ± 0.3	39%
4	*n*-pentyl	1.5 ± 0.1	8%
5	*n*-heptyl	0.7 ± 0.1	ND
6	iPr	1.9[b]	29%
7	iBu	0.4 ± 0.1	3%
8	Ph	1.8 ± 0.1	31%
9	cyclopentyl	1.5 ± 0.1	13%
10	cyclohexyl	ND	ND
11	-$CH_2NH_3^+$ (Gly)	ND	48%

ND: not detected. [a]Yields of ketones were calculated by comparing the area under the curve of extracted ion chromatograms to an internal standard, ^{15}N Arg, and normalized to the native product (entry 2). [b]No duplicate of this reaction run.

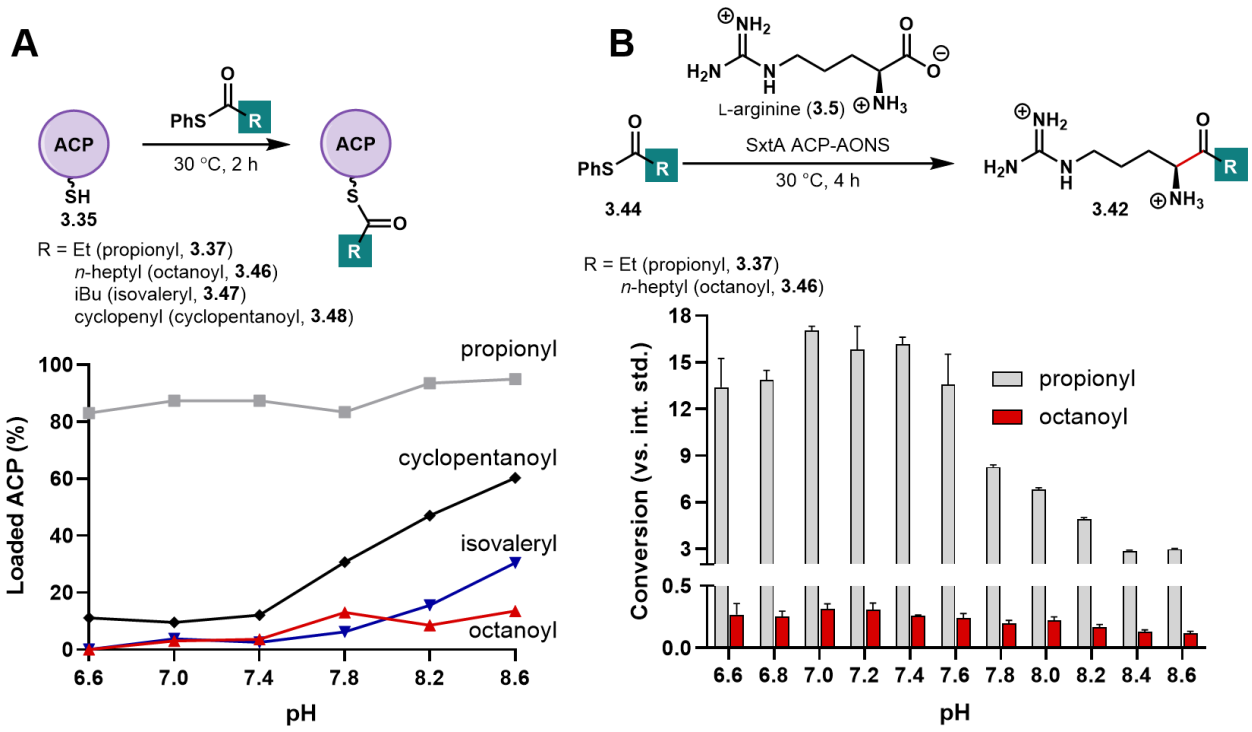


Figure 3.9. Activity of SxtA domains between pH 6.8 and 8.8.
(A) Loading of four acyl groups onto *holo*-ACP. (B) Loading and condensation pH screen with the ACP-AONS didomain. Conditions: 1 mM acyl-SPh, 4 mM Arg, and 10 µM *C. raciborskii* ACP-AONS.

acyl-ACP generally decreased as the size and hydrophobicity of the acyl group increased. The native group, propionyl, had very high loading (89%), while the largest substrate, octanoyl (entry 5), was loaded in negligible amounts to provide very little active acyl-ACP. The glycyl group (entry 11) had modest loading, but showed no ketone formation, indicating that the AONS also does not appear to process positively charged substrates. Adding methanol or THF cosolvents (5% v/v) to increase the solubility of the larger substrates also only suppressed the condensation reaction.

Because synthetic transthioesterification reactions typically occur at more basic pH, we then screened reactions from a range of pH 6.6-8.6 to possibly generate more of the active acyl-ACP species *in situ* (Figure 3.9). A higher pH generally did lead to increase the level of loading (by about 10% for propionyl- and octanoyl-SPh, and over 30% for cyclopentanoyl and isovaleryl-SPh,

see Figure 3.9A), the ketone yields simultaneously decreased (Figure 3.9B). Thus, SxtA ACP more readily undergoes transthioesterification at basic pH, most likely due to the shift in the thiol/thiolate equilibrium, but the AONS domain becomes less active. The best balance of maximum loading and AONS activity for the native propionyl group was at neutral pH 7.0.

From these data, it appeared that only propionyl thioesters that form the native arginine ethyl ketone product **3.8** are feasible for a preparative-scale reaction. When tracking the ketone formation, however, no additional product was converted after 4 h (Figure 3.10). Adding extra equivalents of *holo*-ACP, AONS, propionyl-SPh and L-Arg to find the limiting reagent showed that the AONS was responsible for the stalled reaction. Thus, our current AONS was not robust enough yet to use for semi-preparative or preparative-scale condensation.

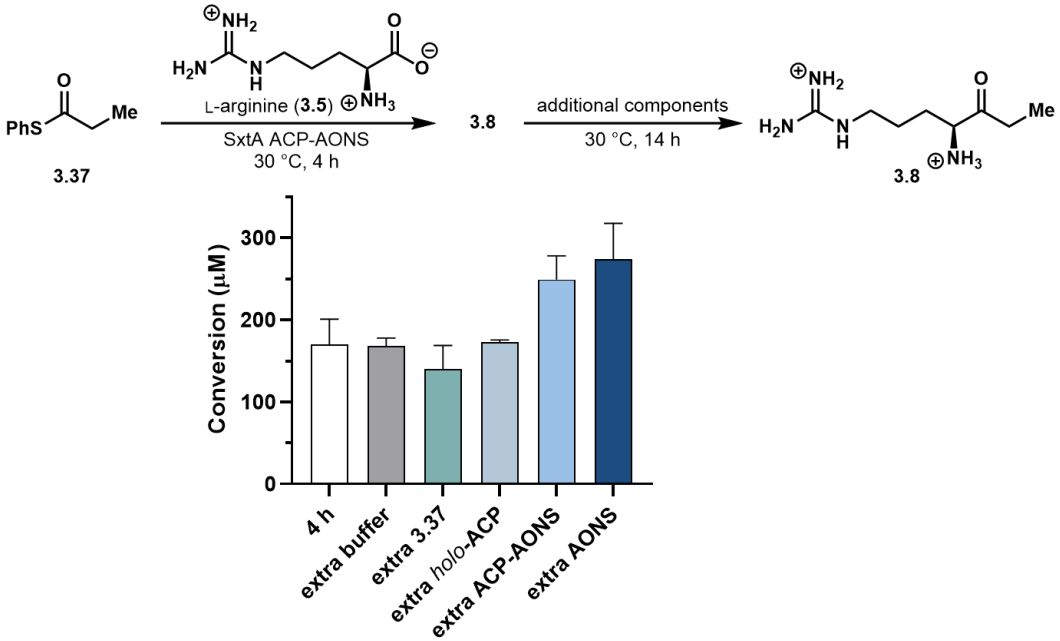


Figure 3.10. Identification of the bottleneck in formation of ketone 3.8.
Conditions: 1 mM propionyl-SPh (**3.37**), 4 mM Arg, 10 μM *C. raciborskii* ACP-AONS, with extra 1 mM propionyl-SPh, 50 μM *holo*-ACP, 10 μM ACP-AONS, or 10 μM ACP(KO)-AONS.

Determination of the product configuration

While the preparative-scale generation experiments of native product ketone **3.8** were paused, we also screened the substrate scopes of three cyanobacterial SxtA AONS homologs for condensation to new ketones (see section 3.3). The homolog from *Microseira wollei* was the most promiscuous with other L-amino acids. Furthermore, when we conducted a presumed negative control condensation reaction with D-Arg and propionyl-pantetheine thioester **3.38**, a small amount of a product matching the mass and retention time of ketone **3.8** was unexpectedly observed (Figure 3.11).

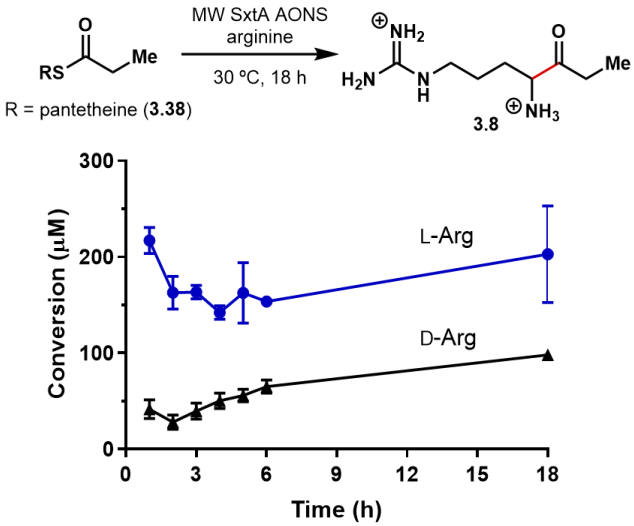


Figure 3.11. Unexpected formation of ketone 3.8 from D-Arg.
Conditions: 1 mM propionyl-pantetheine (**3.38**), 4 mM D- or L-arginine, 10 μM *M. wollei* ACP-AONS.

Although little is known about how amino acid substrates bind in AOS active sites, only one binding mode relative to the PLP cofactor has been observed in AOS-amino acid co-crystal structures, reinforced by a conserved Arg residue (Arg1208 in *M. wollei* SxtA) that interacts with the amino acid carboxylate group.[26,29,58] Additionally, in our homology model of SxtA AONS and published AOS crystal structures, the catalytic Lys residue (*M. wollei* SxtA Lys1083) is located on the *si* face of the PLP cofactor. This only allows an α-proton from an L-substrate in the observed

binding mode and correctly aligned with the PLP p orbitals to be deprotonated (Figure 3.12).[59] The acyl-CoA/-ACP binding site is located on the *re* face of PLP, and addition of the acyl group is proposed to push the carboxylate substituent into the position originally occupied by the alpha proton to promote decarboxylation.[27] The catalytic Lys residue can then reprotonate quinonoid II (3.53) again from the *si* face, affording an L-ketone product. From these data, AOS enzymes have been proposed to only catalyze stereospecific L-ketone formation.

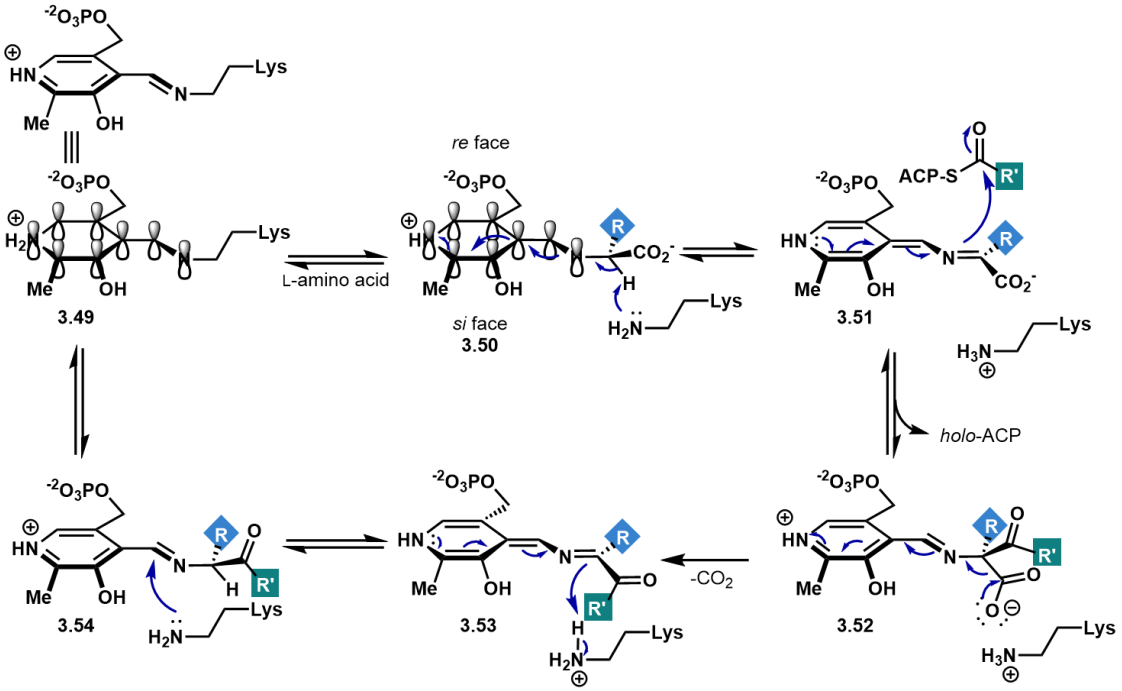


Figure 3.12. Deprotonation and acyl group addition are proposed to occur on opposite faces of the PLP-amino acid complexes in AOS enzymes.

It seems unlikely that a D-Arg substrate is utilizing an alternate binding mode with the α-proton in the same location and swapping the relative orientation of the carboxylate and side chain groups. First, the SxtA AONS residues that interact with a positively charged L-Arg side chain are likely to disfavor interactions with the negatively charged carboxyl group. In addition, the flipped binding mode would place the D-Arg side chain near Arg1208 or in the propionyl group binding tunnel. Again, the positively charged side chain would be repelled by Arg1208, and we never

observed ketone formation with polar or charged acyl groups, indicating that the propionyl binding tunnel probably only allows aliphatic substrates. A second possibility for the initial deprotonation step is via some base on PLP's *re* face in the vicinity of thioester binding site. We could not locate a relevant side chain in our SxtA AONS homology model. In whatever manner the quinonoid and tetra-substituted intermediates are being formed, it is the reprotonation step that sets the product stereocenter. This could be through Lys1083, the thiol byproduct, or another undetermined proton source such as a conserved water molecule.

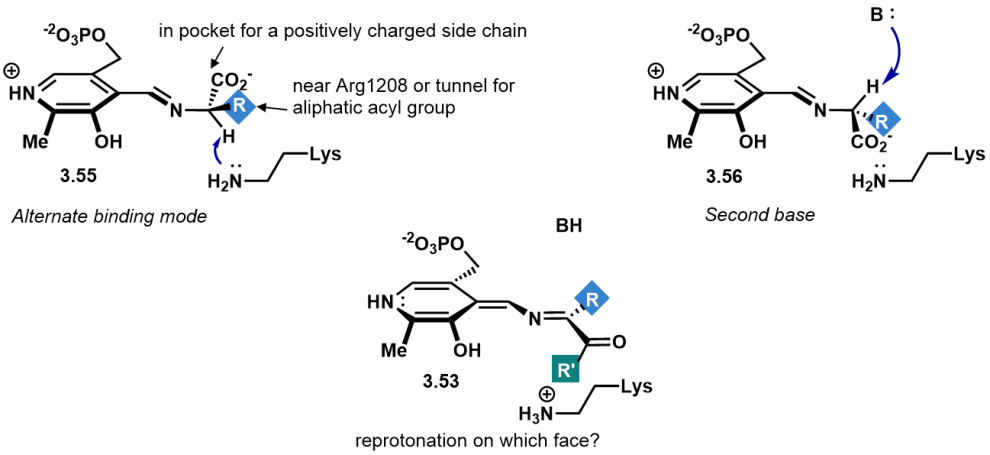


Figure 3.13. Possibilities for deprotonating D-amino acid substrates and reprotonating quinonoid intermediate 3.53.

Of the few AOS proteins that have had their amino acid substrate scopes partially probed, they all use L-configured substrates, or remove the equivalent pro-*R* proton in the case of Gly.[60] Ketone products have also been confirmed with the L-configuration by co-crystallization inside the enzyme active site.[35,61] However, the *Mycobacterium tuberculosis* BioF homolog has been reported to have low levels of activity of D-Ala relative to L-Ala *in vitro*. The final product configuration was not reported.[62] One major challenge to enantiomeric determination is that many α-amino ketones are not configurationally stable in aqueous solution. Indeed, when we incubated our L-synthetic standard of **3.8** overnight in buffered D_2O, we observed incorporation of at least

one deuterium atom by MS in 15% of the population, most likely a mixture of deuteration at both alpha carbons (more detailed discussion in Chapter 4). In order to determine the configuration of the ketones generated from both D- and L-Arg the enantioenrichiment, we would need to trap the product with chiral derivatizing agents as Ploux and coworkers did for BioF to reduce nonenzymatic racemization.[63] An alternative would be reduction of the carbonyl to a hydroxyl group raise the pK_a of the alpha carbons as the sphingosine pathway does.[64] The stereoselectivity of the SxtA AONS reaction cannot be deduced from the structure of the final pathway product, saxitoxin, as we now know that the steps following SxtA catalyzed by SxtG and SxtB erode the alpha stereocenter.[65]

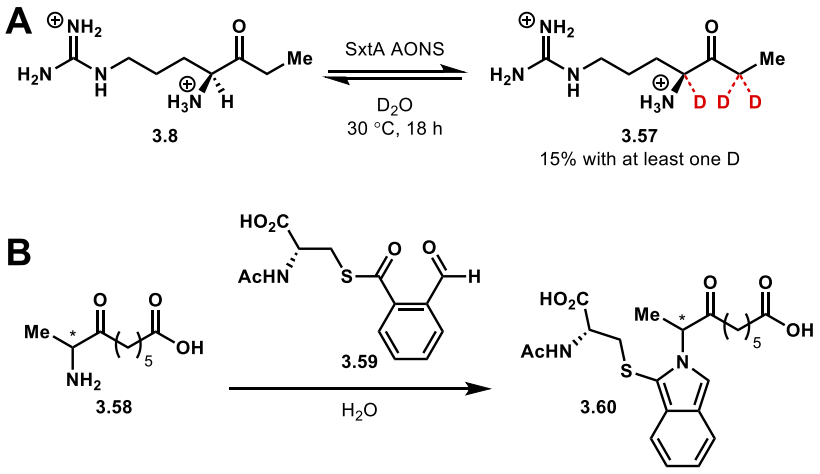


Figure 3.14. Determination of α-amino ketone configurations.
(A) The configuration of the SxtA AONS product ketone **3.8** is not stable in buffered solution. (B) Chiral resolution of the BioF product ketone **3.58** by Ploux and coworkers.[63]

As discussed previously, our SxtA AONS constructs are not robust enough to produce enough ketone product from either arginine enantiomer to analyze the chiral composition. The AONS homology model provides limited information, and we are pursuing a crystal structure with Yongtong Shero Lao in the laboratory of Prof. Janet L. Smith at the University of Michigan. Therefore, investigation of preparative-scale SxtA AONS-catalyzed condensation reactions from arginine was put on hold and we turned our attention toward a directed evolution campaign to

discover more active AONS variants. A more robust enzyme variant hit would produce more of the ketone **3.8** to show any enantioenrichment AONS-catalyzed condensation reactions.

3.3 Condensation to Ketones with Other Side Chains

In Chapter 2 and section 3.2, we demonstrated that SxtA AONS is promiscuous enough to accept a variety of acyl groups and thiol activators. The last possible substrate type to screen was amino acids, which are readily available. However, when *C. raciborskii* SxtA AONS was tested with the remaining nineteen proteinogenic L-amino acids, ornithine and the Arg mimic L-canavanine, only canavanine showed conversion to the corresponding ethyl ketone, as detected by LC-MS.

This substrate scope was disappointingly narrow but can potentially be widened through directed evolution. It is significantly more difficult to start a campaign to evolve for non-native activity when the parent has none, rather than from a low level of activity. Consequently, we obtained the ACP-AONS genes for the other three known cyanobacterial SxtA homologs at the time, *Aphanizomenon gracile*, *Dolichospermum circinale* and *Microseira wollei*, and expressed the proteins.[66,67] A newly reported fifth homolog from *Heteroscytonema crispum*, has been obtained though not expressed.[68]

Some differences with the *C. raciborskii* were immediately visible. While the *C. raciborskii*, *D. circinale*, and *A. gracile* constructs yielded about 20-30 mg of yellow ACP-AONS from 1 L cultures of BAP1 *E. coli*, about 140 mg of *M. wollei* very concentrated orange protein was isolated from the same culture size (Figure 3.15A). *C. raciborskii* protein was still the highest-converting biocatalyst in the native arginine-based reaction (Figure 3.15B), followed by *D. circinale* (DC_AONS), *A. gracile* (AG_AONS) and then *M. wollei* (MW_AONS). In reactions with ten

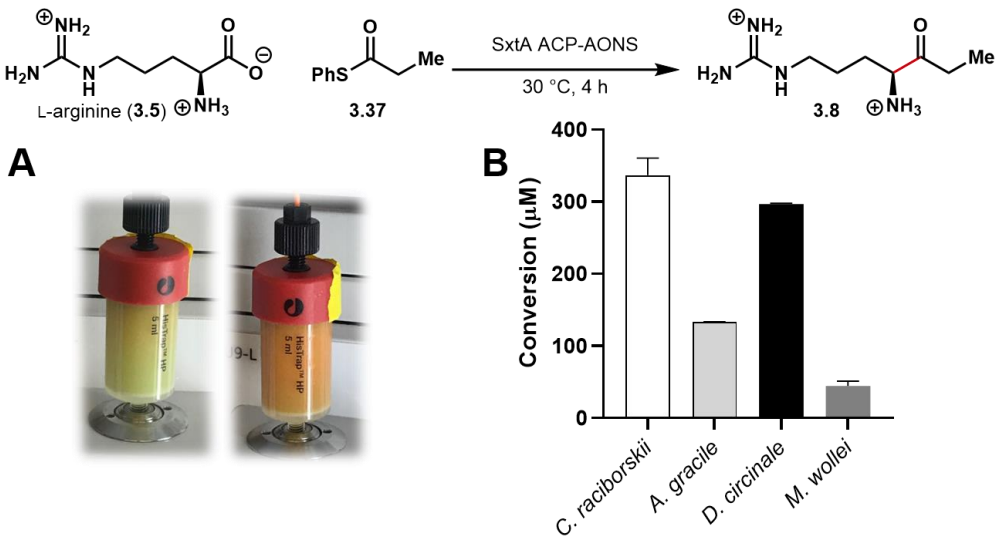


Figure 3.15. Comparison of cyanobacterial SxtA ACP-AONS homologs.
(A) *C. raciborskii* (left) and *M. wollei* (right) protein bound to Ni-NTA resin. (B) Activity of ACP-AONS homologs in formation of ketone **3.8**.

acyl-SPh thioesters, the three homologs converted the same nine substrates as CR_AONS, and one new acyl group, cyclohexyl.

The three new homologs were tested for productive condensation reactions with nineteen proteinogenic L-amino acids and other Arg-related substrates, L-canavanine, L-citrulline, and L-homoarginine (Figure 3.16). CR_AONS and AG_AONS were similar, only affording arginine and canavanine ethyl ketones. In reactions with DC_AONS and MW_AONS, we observed new peaks corresponding to ethyl ketones of Ala, Leu, His, Trp and MW_AONS showed additional possible Ser and Gly products that were not in the negative controls lacking enzyme. We did not observe any decarboxylation or transamination products only. The Ala, Leu and Trp products were verified by LC-MS and MS/MS comparison to synthetic standards prepared by myself and rotation student Jennie Lin. The His, Ser and Gly products have not been confirmed yet. The ketone yields were not quantified, but qualitatively estimated at <1% conversion. Considering the number of natural products and other bioactive compounds containing an indole group, we picked Trp ketone **3.69** as our initial synthetic target for directed evolution (Figure 3.18).

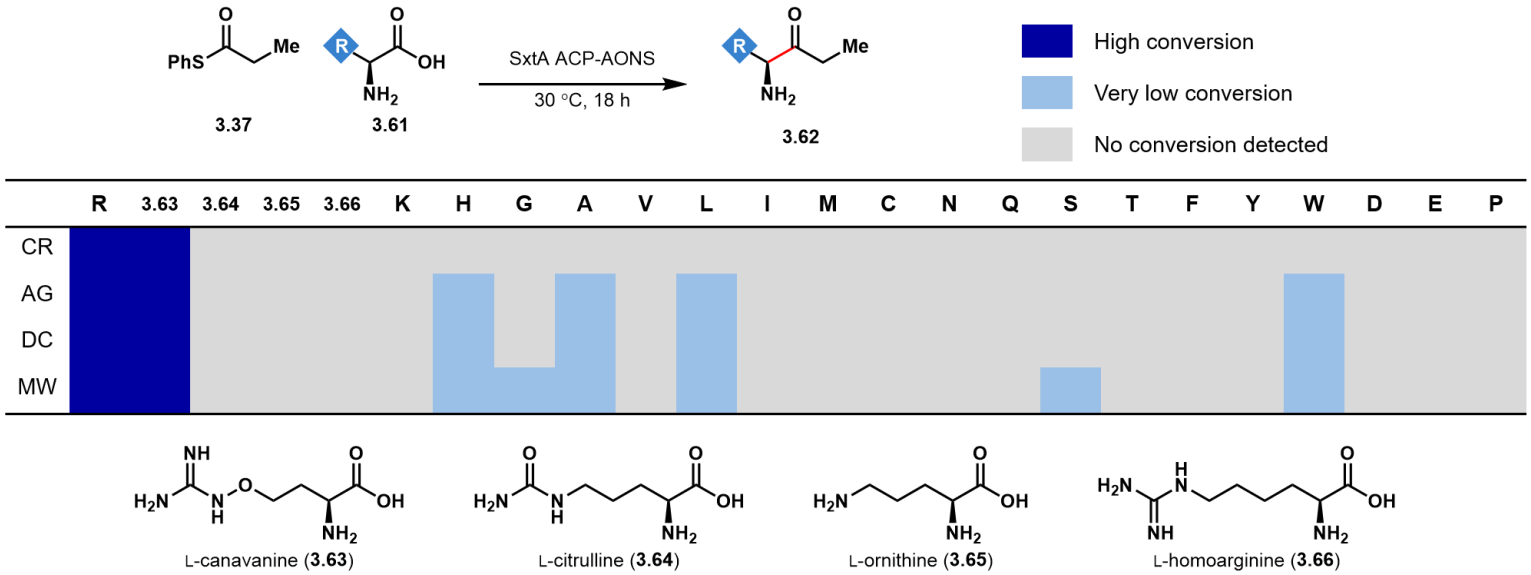


Figure 3.16. Amino acid substrate scope of SxtA ACP-AONS homologs.

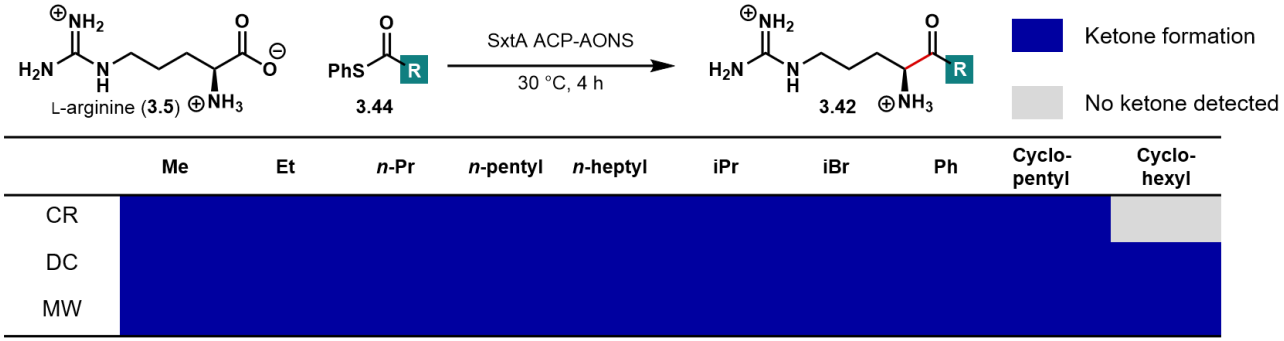


Figure 3.17. Acyl group substrate scope of SxtA ACP-AONS homologs.
The *A. gracile* (AG) homolog was not tested.

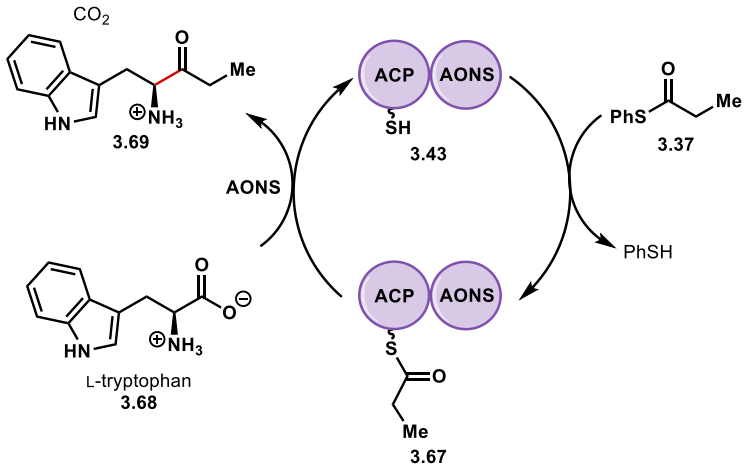


Figure 3.18. Biocatalytic platform for synthesis of tryptophan ethyl ketone **3.69**.

With the highest protein yield and widest amino acid substrate scope, *M. wollei* ACP-AONS was selected as our parent biocatalyst for directed evolution with propionyl-S Ph (**3.37**) as the acyl donor. To avoid most of the lengthy protein purification process, we chose to use a crude format.

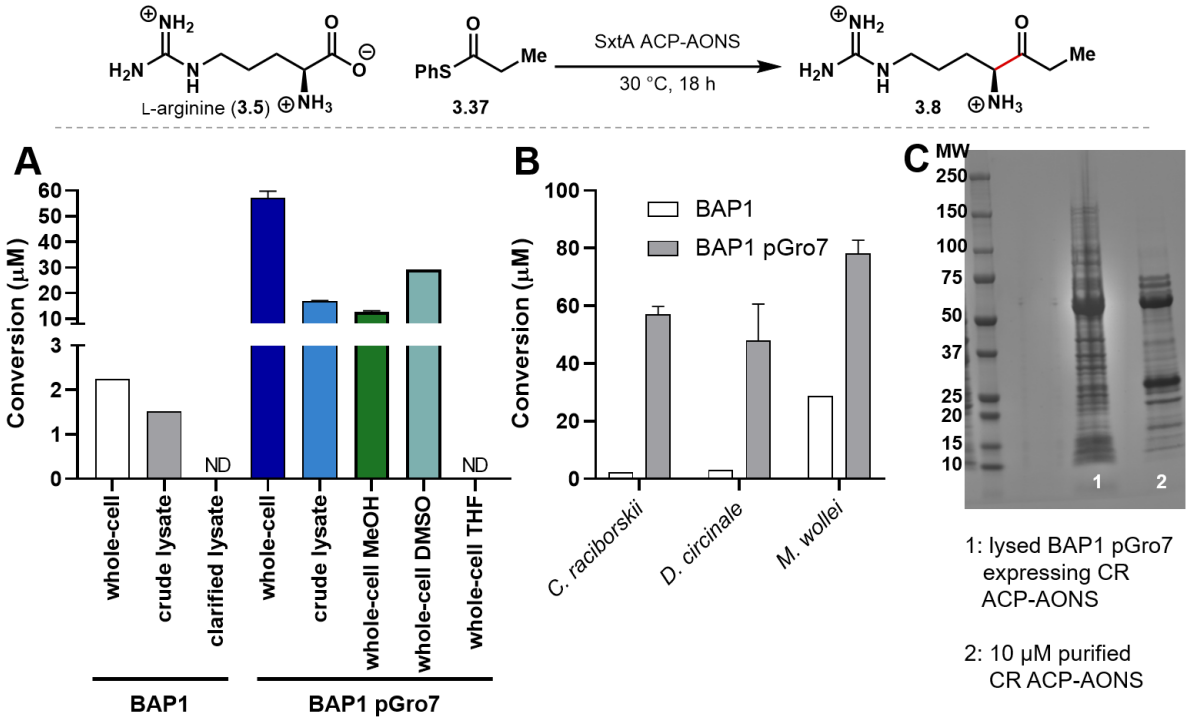


Figure 3.19. Conversion of arginine to the native ketone product in crude formats. (A) With *C. raciborskii*, comparison of whole-cell and lysates with organic solvents in BAP1 and BAP1 pGro7 *E. coli*. (B) Whole-cell activity with SxtA ACP-AONS homologs. (C) An SDS-PAGE gel of crude lysate and purified protein reactions shows that they contain similar catalyst loading (expected molecular weight: 61 kDa).

Before we obtained the new SxtA ACP-AONS homologs, we had compared the relative yields of native product ketone **3.8** for whole-cell, crude lysate, and clarified lysate reactions with *C.raciborskii* ACP-AONS (Figure 3.19A). The less processing, the more active the biocatalyst, so whole cells in a 96-well format were also used for *M. wollei* ACP-AONS. Switching from the base BAP1 *E. coli* strain to one also containing the pGro7 plasmid (Takara Biosciences) that encodes for GroES and GroEL chaperone proteins also increased conversion to **3.69** (Figure 3.19B). As in the *in vitro* purified protein reactions, the addition of organic co-solvents gave no advantage.

With the target, biocatalyst, and reaction format decided, the next challenge was the high-throughput screening of the enzymatic reactions to identify improved variants. Our normal LC-MS method was 12 min in total, processing five samples an hour and a 96-well plate in 20 hours.

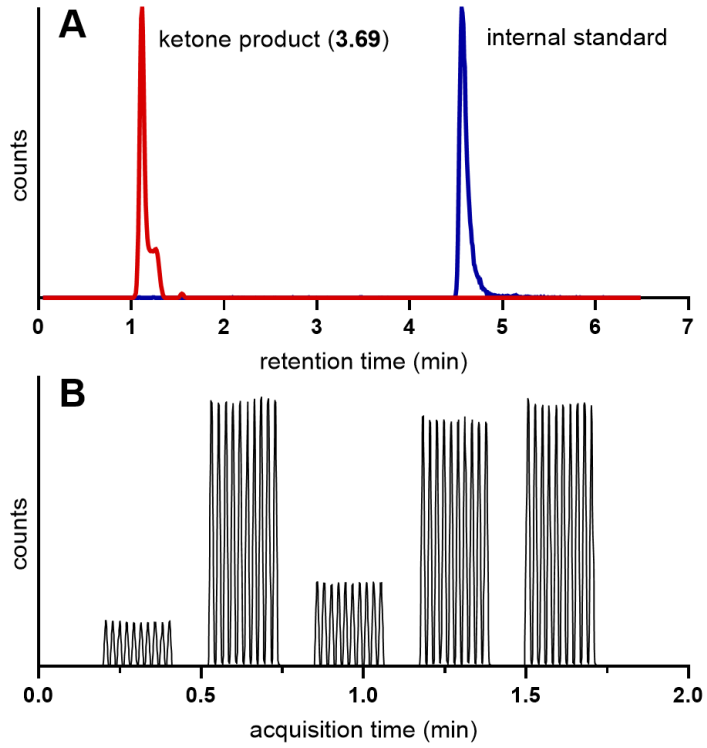


Figure 3.20. Mass spectrometry analysis of the SxtA ACP-AONS random mutagenesis library.
(A) Low-throughput LC-MS (sample identification and quantification by 6 min, total time with re-equilibration: 12 min). (B) Representative high-throughput droplet MS trace, (first droplet readout around 15 s). Each peak represents one droplet.[69] Data courtesy of LeeAnne Wang.

This was far too low throughput for effective analysis, so we began a collaboration with LeeAnne Wang and Emory Payne of Prof. Robert Kennedy's lab at the University of Michigan for rapid screening. Their droplet microfluidic MS platform up to a thousand samples[69,70] an hour, a much higher-throughput than our TOF and Q-TOF LC-MS instruments (Figure 3.20). The Kennedy lab was able to optimize detection of our synthetic standard of ketone **3.69** relative to the buffer and other cell material. Droplet MS is less sensitive than LC-MS at detecting our product among the complex reaction buffer, so any potential hits must be verified by LC-MS in a secondary screen.

In our first round of directed evolution, we employed error-prone PCR to introduce random mutations in the AONS domain. This generated a 1.5 kb "megaprimer" for whole-plasmid PCR with wild-type ACP-AONS using the MEGAWHOP[71] method. After digestion of the parent plasmid, ligation and transformation into BAP1 pGro7 *E. coli*, we picked colonies for ten 96-well plates (Figure 3.21). Sanger sequencing of a few random clones revealed 1-3 mutations per library member.

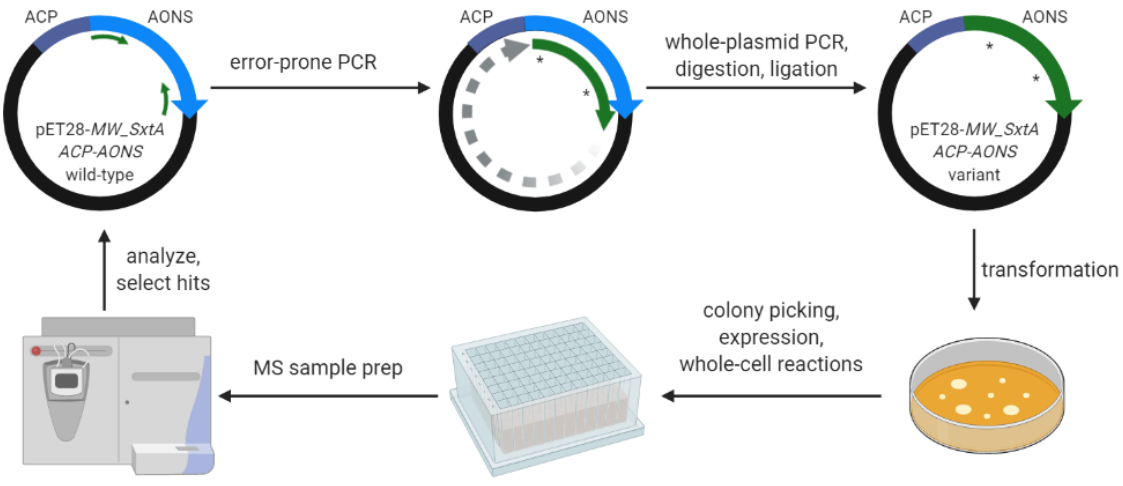


Figure 3.21. Library generation and high-throughput screening for directed evolution of SxtA ACP-AONS.
This image was created on Biorender.com.

We conducted whole-cell reactions with the first-generation library and prepared sample plates for the Kennedy group for analysis. Even in the positive wild-type control wells, the desired Trp ketone **3.69** was difficult to detect relative to the empty plasmid negative controls, with SxtA ACP-AONS preferring to produce its native ketone **3.8** from endogenous Arg instead. Nevertheless, we identified three wells with potentially slightly increased Trp ketone formation activity. Sanger sequencing for two of the hits showed that the mutations (F979L and I1076T using the whole module numbering) were located at seemingly random locations in the AONS domain, distal from the active site. A second-generation library from these two sequences has not been generated.

Because the conversions from these new variants were barely more than those of the wild-type enzyme, we temporarily put this particular project on hold. In the meantime, Sarah Ackenhusen in our lab has designed a new site-saturation mutagenesis (SSM) library for the standalone AONS domain, targeting 96 sites near the active site. These libraries will be screened first for their ability to catalyze formation of tetra-substituted amino acid derivatives similar to **3.15** (see Figure 3.3). They can also be adapted to the Trp ketone condensation reaction with some modification.

At the equivalent of 1% conversion (10 μM) of the synthetic ketone **3.69** standard relative to the reaction matrix, the signal-to-noise ratio was good, so future adjustments would begin by optimizing detection for even lower conversions. Adjustments to improve the low signal in the next rounds of screening include using nano-electrospray ionization for higher sensitivity[72] or employing solid phase extraction (SPE) cartridges prior to the MS. SPE cleanup is faster than liquid chromatography separation and would remove many of the competing buffer components. This process is possible in-line with the MS on a system such as the Agilent RapidFire 365 in the University of Michigan's Chemistry Department.

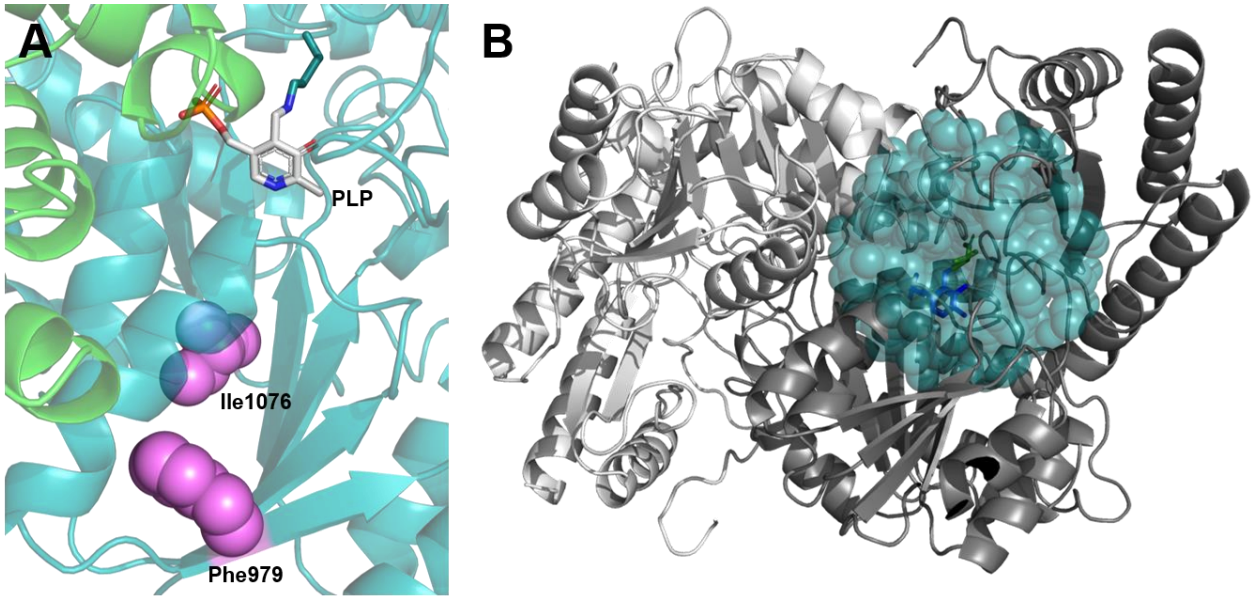


Figure 3.22. Homology models of *M. wollei* SxtA AONS.
All AOS members have been found to be homodimers. The AONS homology model was generated by the Phyre2 server (based on PDB 1DJ9, 27% identity with SxtA AONS).[29] ACPs tend to be very flexible and were not modeled. (A) Close-up of the AONS active site, which lies near the boundary of the two monomers. The lysine-bound PLP cofactor of the teal monomer is shown in gray; the second monomer is in green. The two mutated residues found thus far are highlighted as magenta spheres. (B) The 96 residues within 12 Å of the active site chosen for site-saturation mutagenesis are shown as teal spheres. Image by Sarah Ackenhusen.

Due to the small volumes of droplet microfluidics (e.g. 45 nL per droplet vs. 0.5 µL for LC-MS after a 100x dilution from a 250 µL whole-cell reaction), our enzymatic reactions can similarly be miniaturized further. With recent advances in reagent addition techniques and other reaction methods[73] at droplet scales, it may be possible to run the whole-cell reactions in droplets as well. Using pantetheine thioester **3.38** becomes more economically feasible at a smaller scale, only requiring a few microliters of the acyl donor for the entire experiment instead of microliters per well. Additionally, removing the ACP domain decreases the protein weight by 10 kDa and we would no longer need to ensure post-translational modification (e.g., the Ppant cofactor), further decreasing the cost per reaction.

3.4 Conclusions

As the AONS domain of SxtA domain modifies α-amino acids to α-amino ketones in a single protecting group-free step, we saw the potential for this enzyme to serve as a general amino acid-modifying biocatalyst. This chapter summarizes our efforts to chemoenzymatically synthesize α-amino ketones, catalyzed by standalone SxtA AONS or AONS with its associated ACP domain functioning in tandem as a didomain.

After extensive screening of the acyl group and activating thiol substrate scope SxtA AONS, we optimized the reaction conditions using the ACP-AONS didomain and inexpensive thioester donors to generate the active acyl-ACP species *in situ*. We intended to run preparative-scale experiments that synthesize larger amounts of non-native arginine ketone products, which could potentially serve as precursors for new saxitoxin derivatives. The native reaction to produce the native ketone **3.8** had low conversion, and the generation of non-native ketones was even lower. When we later observed that the D-enantiomer of the AONS substrate L-Arg was unexpectedly also converted to arginine ethyl ketone, we considered utilizing the same platform to study the stereoselectivity of SxtA AONS-catalyzed condensation. The developed our biocatalytic platform lowers the cost of these chemoenzymatic α-amino ketone-forming reactions, and efforts are current aimed at improving the activity of SxtA AONS activity.

Furthermore, we screened the amino acid substrate scopes of our original SxtA ACP-AONS homolog from *C. raciborskii* and three other cyanobacterial homologs. *M. wollei* ACP-AONS demonstrated that it was likely also synthesizing ethyl ketones with seven other side chains in addition to its native L-arginine substrate. In collaboration with the Kennedy group, we commenced a directed evolution campaign to evolve variants that are more active toward one non-native substrate, L-tryptophan. Two potential variants have been identified in the first generation

after random mutagenesis and high-throughput droplet MS screening. An SSM library of SxtA AONS instead.

In our optimized reaction conditions, we fulfilled our original goal of using ACPs at a catalytic concentration. However, the main barrier to using AOS enzymes as α-amino ketone-forming biocatalysts in this chapter were the limited turnovers of our chosen PE, SxtA AONS. In future efforts, we will investigate the potential causes of the AONS' degrading activity in order to focus directed evolution campaigns on improving overall enzyme stability and robustness first, before expanding the substrate scopes further.

As *holo*-ACPs are still expensive to produce, either using the slower-growing BAP1 strain for endogenous post-translational modification or enzymatic modification *in vitro* with CoASH, we also have the opportunity to engineer SxtA AONS to be more active with acyl substrates free in solution. Thioester sources such as thiophenols would be much more cost-effective. These projects will be further investigated by other lab members. We anticipate that with more engineering, SxtA AONS will be a useful tool for the enzymatic transformation of any combination of α-amino acid and thioester to α-amino ketones.

3.5 Experimental

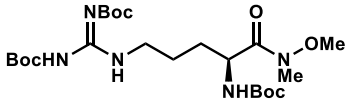
I. Chemical synthesis

A. General information

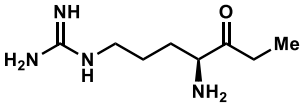
All reagents were used as received unless otherwise noted. Reactions were carried out under a nitrogen atmosphere using standard Schlenck techniques unless otherwise noted. Solvents were degassed and dried over aluminum columns on an MBraun solvent system (Inert Corporation, Model PS-00-3). Trifluoroacetic acid was distilled prior to use. Reactions were monitored by thin layer chromatography using Machery-Nagel 60 F_{254} precoated silica TLC plates (0.25 mm) or Merck Silica Gel 60 RP-18 WF-254S precoated silica TLC plates (0.25 mm) which were visualized using UV, ninhydrin, *p*-anisaldehyde, CAM, DNP, or bromocresol green stain. Flash column chromatography was performed using Machery-Nagel 60 μm (230-400 mesh) silica gel. All compounds purified by column chromatography were sufficiently pure for use in further experiments unless otherwise indicated. ^{1}H and ^{13}C NMR spectra were obtained in $CDCl_3$ or CD_3OD at rt (25 °C), unless otherwise noted, on Varian 400 MHz, Varian 500 MHz or Varian 600 MHz spectrometers. Chemical shifts of ^{1}H NMR spectra were recorded in parts per million (ppm) on the δ scale. High resolution electrospray mass spectra were obtained on an Agilent G6545A quadrupole-time of flight mass spectrometer in positive mode with an Agilent 1290 UPLC system. Solvent A = water with 0.1% formic acid. Solvent B = 95% acetonitrile, 5% water and 0.1% formic acid. IR spectra were recorded on a Perkin-Elmer Spectrum BX FT-IR spectrometer. Optical rotations were measured at rt in CH_3OH, unless otherwise noted, on a Jasco P-2000 polarimeter.

The synthesis and characterization of the arginine ketone standards was done by Dr. Meagan Hinze. Leucine and alanine ketones were prepared by Jennie Lin.

B. Compound synthesis

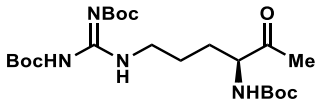


3.S1



3.8

Tri-Boc-protected arginine Weinreb amide (**3.S1**) and arginine ethyl ketone (**3.8**) were synthesized according to the procedure described by Tsuchiya et al.[25]

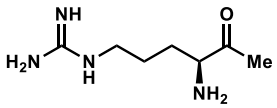


3.S2

Tri-Boc-protected arginine methyl ketone (3.S2):

To a solution of tri-Boc-protected arginine Weinreb amide (**3.S1**, 72.0 mg, 0.139 mmol, 1.0 equiv) dissolved in THF (6.9 mL) was added a 1 M solution of zinc chloride in diethyl ether (13.9 μL, 0.013 mmol, 0.1 equiv) at 0 °C. After five min, a 3 M solution of methyl magnesium bromide in diethyl ether (695 μL, 2.09 mmol, 15.0 equiv) was added at 0 °C. The reaction was allowed to warm to room temperature and stirred for 14 h. The reaction was poured into 21 mL of 0.1 M HCl. The acidified aqueous phase was extracted with EtOAc (3x). The organic layers were combined, washed with brine, dried over Na_2SO_4, filtered, and evaporated under reduced pressure to afford an oil. The crude residue was purified by flash column chromatography (2:1 hexanes/EtOAc v/v) to afford 28.5 mg of the title compound (43% yield) as a white foam. R_f = 0.52 (2:1 hexanes/EtOAc v/v); **^{1}H NMR** (300 MHz, $CDCl_3$) δ 11.46 (s, 1H), 8.33 (t, *J* = 5.4, 1H), 5.36 (d, *J* = 7.6, 1H), 4.35-4.28 (m, 1H), 3.42 (q, *J* = 6.5, 2H), 2.20 (s, 3H), 1.95-1.85 (m, 1H), 1.66-1.52 (m, 3H), 1.48 (s, 9H), 1.47 (2, 9H), 1.42 (s, 9H); **^{13}C NMR** (150 MHz, $CDCl_3$) δ 207.1, 163.6, 156.4, 155.6, 153.4, 83.3, 79.9, 79.4, 59.7, 40.3, 28.5, 28.4, 28.2, 27.1, 25.2; **IR** (thin film, cm^{-1}) 3329, 2978, 1715,

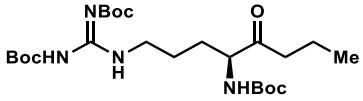
1614, 1498, 1414; **HRMS** (ESI) *m/z* calculated for $C_{22}H_{41}N_4O_7^+$ $[M+H]^+$ 473.2970, found 473.2977.



3.9

(*S*)-1-(4-amino-5-oxohexyl)guanidine (3.9):

To a solution of tri-Boc protected arginine methyl ketone (**3.S2**, 28.5 mg, 0.060 mmol, 1.0 equiv) dissolved in DCM (1.4 mL) was added trifluoroacetic acid (TFA, 1.4 mL) at rt. The reaction was stirred for 3 h. After addition of toluene (2.0 mL), the solution was evaporated under reduced pressure to afford 24.0 mg of the title compound as a bis-TFA salt with minor impurities (99% yield) as a brown foam. R_f = 0.10 (RP_{18} TLC, 2:1 MeCN/H_2O v/v); $[\alpha]_D^{24}$ = 0.310 (*c* = 0.014 g/mL, MeOH); **^{1}H NMR** (600 MHz, CD_3OD) δ 4.22-4.17 (m, 1H), 3.24 t, *J* = 6.3, 2H), 2.28 (s, 3H), 2.08 (t, *J* = 13.0, 1H), 1.91-1.81 (m, 1H), 1.76-1.67 (m, 1H), 1.67-1.57 (m, 1H); **^{13}C NMR** (150 MHz, CD_3OD) δ 204.2, 158.7, 60.0, 41.7, 27.5, 26.4, 25.2; **IR** (thin film, cm^{-1}) 3363, 3187, 1670, 1199, 1133; **HRMS** (ESI) *m/z* calculated for $C_7H_{17}N_4O^+$ $[M+H]^+$ 173.1397, found 173.1398.

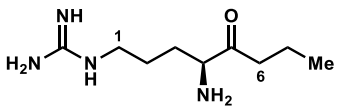


3.S3

Tri-Boc protected arginine propyl ketone (3.S3):

To a solution of tri-Boc-protected arginine Weinreb amide (**3.S1**, 50.0 mg, 0.0966 mmol, 1.0 equiv) dissolved in THF (4.8 mL) was added a 2 M solution of propyl magnesium chloride in diethyl ether (725 µL, 1.45 mmol, 15.0 equiv) at 0 °C. The reaction was allowed to warm to rt and stirred for 3.25 h. The reaction was poured into 15 mL of 0.1 M HCl. The acidified aqueous phase

was extracted with EtOAc (3x). The organic layers were combined, washed with brine, dried over Na_2SO_4, filtered, and evaporated under reduced pressure to afford an oil. The crude residue was purified by flash column chromatography (4:1 hexanes/EtOAc v/v) to afford 30.0 mg of the title compound (62% yield) as a colorless oil. R_f = 0.52 (2:1 hexanes/EtOAc v/v); **^{1}H NMR** (600 MHz, $CDCl_3$) δ 11.47 (s, 1H), 8.32 (d, *J* = 5.6, 1H), 5.34 (d, *J* = 7.8, 1H), 4.31 (dt, *J* = 11.7, 5.5, 1H), 3.41 (qd, *J* = 7.1, 3.0, 2H), 2.46 (q, *J* = 6.8, 6.4, 2H), 1.89 (d, *J* =15.8, 1H), 1.61 (h, *J* =7.2, 3H), 1.57-1.50 (m, 2H), 1.48 (s, 9H), 1.48 (s, 9H), 1.42 (s, 9H), 0.90 (t, *J* = 7.4, 3H); **^{13}C NMR** (150 MHz, $CDCl_3$) δ 209.3, 163.7, 156.3, 155.6, 153.4, 83.3, 79.9, 79.4, 59.0, 41.7, 40.4, 28.8, 28.5, 28.4, 28.2, 25.1, 17.1, 13.9; **IR** (thin film, cm^{-1}) 3331, 2976, 2933, 1714, 1637, 1614, 1573, 1454, 1413; **HRMS** (ESI) *m/z* calculated for $C_{24}H_{45}N_4O_7^+$ $[M+H]^+$ 501.3283, found 501.3290.



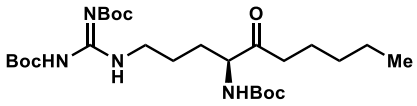
3.42

(*S*)-1-(4-amino-5-oxooctyl)guanidine (3.42a):

To a solution of tri-Boc protected arginine propyl ketone (**3.S3**, 26.3 mg, 0.0525 mmol, 1.0 equiv) dissolved in DCM (1.2 mL) was added trifluoroacetic acid (TFA, 1.2 mL) at rt. The reaction was stirred for 4.3 h. After addition of toluene (2.0 mL), the solution was evaporated under reduced pressure to afford 20.9 mg of the title compound as a bis-TFA salt with minor impurities (93% yield) as a colorless oil. R_f = 0.23 (RP_{18} TLC, 3:1 0.1% formic acid in H_2O/MeOH v/v); **^{1}H NMR** (600 MHz, CD_3OD) δ 4.18 (dd, *J* = 7.8, 4.2, 1H), 3.25 (td, *J* = 6.9, 3.0, 2H), 2.67-2.54 (m, 2H), 2.11-2.04 (m, 1H), 1.86 (dddd, *J* = 14.6, 12.1, 7.7, 4.6, 1H), 1.77-1.69 (m, 1H), 1.69-1.60 (m, 3H), 0.96 (t, *J* = 7.4, 3H); **^{13}C NMR** (150 MHz, CD_3OD)* δ 206.6, 158.8, 59.5, 41.7, 27.8, 25.3, 17.7,

13.8; **IR** (thin film, cm^{-1}) 3355, 3177, 2938, 1669, 1558, 1539; **HRMS** (ESI) *m/z* calculated for $C_9H_{21}N_4O^+$ $[M+H]^+$ 201.1710, found 201.1712.

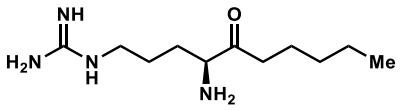
* The respective ^{13}C signals of C-1 and C-6 coalesce as verified by HSQC.



3.S4

Tri-Boc protected arginine pentyl ketone (3.S4):

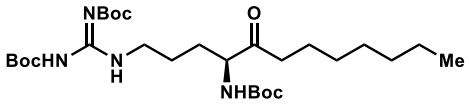
To a solution of tri-Boc-protected arginine Weinreb amide (**3.S1**, 50.0 mg, 0.0966 mmol, 1.0 equiv) dissolved in THF (4.8 mL) was added a 2 M solution of pentyl magnesium bromide in diethyl ether (725 µL, 1.45 mmol, 15.0 equiv) at 0 °C. The reaction was allowed to warm to rt and stirred for 14.2 h. The reaction was poured into 15 mL of 0.1 M HCl. The acidified aqueous phase was extracted with EtOAc (3x). The organic layers were combined, washed with brine, dried over Na_2SO_4, filtered, and evaporated under reduced pressure to afford an oil. The crude residue was purified by flash column chromatography (4:1 hexanes/EtOAc v/v) to afford 8.0 mg of the title compound (16% yield) as a colorless oil. R_f = 0.60 (2:1 hexanes/EtOAc v/v); **^{1}H NMR** (400 MHz, $CDCl_3$) δ 11.48 (s, 1H), 8.33 (t, *J* = 5.0, 1H), 5.34 (d, *J* = 8.1, 1H), 4.32 (s, 1H), 3.42 (dd, *J* = 9.8, 4.5, 2H), 2.48 (td, *J* = 7.3, 2.6, 2H), 1.89 (br s, 1H), 1.62-1.56 (m, 3H), 1.50 (s, 9H), 1.49 (s, 9H), 1.44 (s, 9H), 1.36-1.22 (m, 6H), 0.88 (t, *J* = 7.0, 3H); **^{13}C NMR** (100 MHz, $CDCl_3$) δ 209.4, 163.7, 156.4, 155.6, 153.5, 83.3, 79.9, 79.4, 59.0, 40.4, 39.8, 31.5, 28.9, 28.5, 28.4, 28.2, 25.1, 23.4, 22.6, 14.0; **IR** (thin film, cm^{-1}) 3331, 2977, 2931, 1717, 1639, 1616, 1573, 1455, 1414; **HRMS** (ESI) *m/z* calculated for $C_{26}H_{49}N_4O_7^+$ $[M+H]^+$ 529.3596, found 529.3602.



3.42b

(*S*)-1-(4-amino-5-oxodecyl)guanidine (3.42b):

To a solution of tri-Boc protected arginine pentyl ketone (**3.S4**, 7.5 mg, 0.014 mmol, 1.0 equiv) dissolved in DCM (0.4 mL) was added trifluoroacetic acid (TFA, 0.4 mL) at rt. The reaction was stirred for 4.6 h. After addition of toluene (2.0 mL), the solution was evaporated under reduced pressure to afford 3.9 mg of the title compound as a bis-TFA salt with minor impurities (60% yield) as a colorless oil. R_f = 0.15 (RP_{18} TLC, 3:1 0.1% formic acid in H_2O/MeOH v/v); **^{1}H NMR** (600 MHz, CD_3OD) δ 4.17 (dd, *J* = 7.7, 4.2, 1H), 3.24 (td, *J* = 6.9, 3.2, 2H), 2.61 (qt, *J* = 17.8, 7.3, 2H), 2.10-2.02 (m, 1H), 1.85 (dddd, *J* = 14.7, 12.2, 7.7, 4.7, 1H), 1.75-1.67 (m, 1H), 1.65-1.57 (m, 3H), 1.37-1.31 (m, 4H), 0.91 (t, *J* = 7.1, 3H); **^{13}C NMR** (150 MHz, $CDCl_3$) δ 205.3, 157.3, 58.1, 40.3, 38.4, 30.9, 26.4, 23.9, 22.5, 22.1, 12.8; **IR** (thin film, cm^{-1}) 3171, 2932, 2874, 1670, 1533,1432; **HRMS** (ESI) *m/z* calculated for $C_{11}H_{25}N_4O^+$ $[M+H]^+$ 229.2023, found 229.2025.

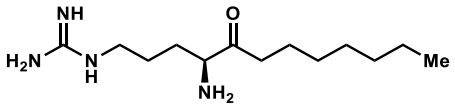


3.S5

Tri-Boc protected arginine heptyl ketone (3.S5):

A solution of heptyl magnesium bromide was prepared by the following procedure. To a suspension of magnesium turnings (148 mg, 6.11 mmol, 1.2 equiv) and iodine (1.0 mg, 0.0078 mmol, 0.001 equiv) in diethyl ether (4.0 mL) was added 1-bromoheptane in 100-200 μL portions over 45 min at rt until all 1-bromoheptane (800 μL, 5.09 mmol, 1.0 equiv) had been added. After 3.75 h, the solution was transferred via cannula to a Schlenk flask and determined to be approximately 0.5 M by colorimetric titration using the method described by Paquette et al.[74]

To a solution of tri-Boc-protected arginine Weinreb amide (**3.S1**, 50.0 mg, 0.0966 mmol, 1.0 equiv) dissolved in THF (4.8 mL) was added a 0.5 M solution of heptyl magnesium bromide in diethyl ether (2.90 mL, 1.45 mmol, 15.0 equiv) at 0 °C. The ice bath was removed after 10 min, and the reaction flask allowed to warm to rt and stirred for 12.4 h. The reaction was poured into 15 mL of 0.1 M HCl. The acidified aqueous phase was extracted with EtOAc (3x). The organic layers were combined, washed with brine, dried over Na_2SO_4, filtered, and evaporated under reduced pressure to afford an oil. The crude residue was purified by flash column chromatography and required two purifications under different conditions (4:1 hexanes/EtOAc and 95:5 DCM/ether v/v) to afford 11.3 mg of the title compound (21% yield) as a colorless oil. R_f = 0.58 (2:1 hexanes/EtOAc v/v), R_f = 0.65 (9:1 DCM/ether v/v); **^{1}H NMR** (500 MHz, $CDCl_3$) δ 11.48 (s, 1H), 8.33 (t, *J* = 5.3, 1H), 5.34 (d, *J* = 7.8, 1H), 4.32 (dt, *J* = 11.6, 5.6, 1H), 3.42 (dh, *J* = 13.3, 7.2, 2H), 2.48 (td, *J* = 7.4, 3.1, 2H), 1.94-1.83 (m, 1H), 1.66-1.51 (m, 5H), 1.49 (s, 9H), 1.48 (s, 9H), 1.43 (s, 9H), 1.32-1.23 (m, 8H), 0.87 (t, *J* = 6.7, 3H); **^{13}C NMR** (125 MHz, $CDCl_3$) δ 209.4, 163.7, 156.3, 155.6, 153.4, 83.3, 79.9, 79.4, 59.0, 40.4, 39.9, 31.8, 29.3, 29.2, 28.9, 28.5, 28.4, 28.2, 25.1, 23.7, 22.7, 14.2; **IR** (thin film, cm^{-1}) 3333, 2977, 2929, 2856, 1717, 1639, 1616, 1573, 1502, 1454, 1415; **HRMS** (ESI) *m/z* calculated for $C_{28}H_{53}N_4O_7^+$ $[M+H]^+$ 557.3909, found 557.3907.

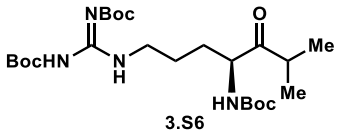


3.42c

(*S*)-1-(4-amino-5-oxododecyl)guanidine (3.42c):

To a solution of tri-Boc protected arginine heptyl ketone (**3.S5**, 11.3 mg, 0.0202 mmol, 1.0 equiv) dissolved in DCM (0.5 mL) was added trifluoroacetic acid (TFA, 0.5 mL) at rt. The reaction was stirred for 3.2 h. After addition of dichloromethane (2.0 mL), the solution was evaporated under

reduced pressure to afford 9.8 mg of the title compound as a bis-TFA salt (quant) as a colorless oil. R_f = 0.15 (RP_{18} TLC, 3:1 0.1% formic acid in H_2O/MeOH v/v); **^{1}H NMR** (500 MHz, CD_3OD) δ 4.19 (dd, *J* = 7.7, 4.2, 1H), 3.24 (td, *J* = 6.9, 1.8, 2H), 2.71-2.54 (m, 2H), 2.07 (ddt, *J* = 14.4, 11.8, 4.6, 1H), 1.90-1.81 (m, 1H), 1.72 (tq, *J* = 12.2, 6.6, 1H), 1.68-1.56 (m, 3H), 1.36-1.28 (m, 8H), 0.90 (t, *J* = 6.7, 3H); **^{13}C NMR** (125 MHz, CD_3OD) δ 206.8, 158.8, 59.5, 41.7, 39.9, 32.9, 30.2, 30.1, 27.9, 25.3, 24.3, 23.7, 14.4; **IR** (thin film, cm^{-1}) 3353, 3179, 2956, 2929, 2858, 1670, 1533, 1456, 1433; **HRMS** (ESI) *m/z* calculated for $C_{13}H_{29}N_4O^+$ $[M+H]^+$ 257.2336, found 257.2338.

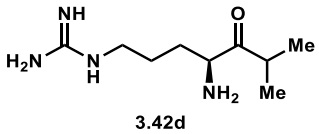


3.S6

Tri-Boc protected arginine isopropyl ketone (3.S6):

To a solution of tri-Boc-protected arginine Weinreb amide (**3.S1**, 50.0 mg, 0.0966 mmol, 1.0 equiv) dissolved in THF (1.9 mL) was added a 2 M solution of isopropyl magnesium chloride in tetrahydrofuran (725 μL, 1.45 mmol, 15.0 equiv) at 0 °C. The reaction was heated at 45 °C and stirred for 3 h. The reaction was poured into 15 mL of 0.1 M HCl. The acidified aqueous phase was extracted with EtOAc (3x). The organic layers were combined, washed with brine, dried over Na_2SO_4, filtered, and evaporated under reduced pressure to afford an oil. The crude residue was purified by flash column chromatography (4:1 hexanes/EtOAc v/v) to afford 6.1 mg of the title compound (13% yield) as a colorless oil. R_f = 0.61 (2:1 hexanes/EtOAc v/v); **^{1}H NMR** (600 MHz, $CDCl_3$) δ 11.47 (s, 1H), 8.33 (t, *J* = 5.5, 1H), 5.33-5.28 (m, 1H), 4.50 (t, *J* = 6.7, 1H), 3.42 (q, *J* = 6.5, 2H), 2.82 (p, *J* = 6.8, 1H), 1.92-1.85 (m, 1H), 1.59-1.50 (m, 3H), 1.49 (s, 9H), 1.49 (s, 9H), 1.44 (s, 9H), 1.13 (d, *J* = 7.0, 3H), 1.09 (d, *J* = 6.7, 3H); **^{13}C NMR** (150 MHz, $CDCl_3$) δ 213.0,

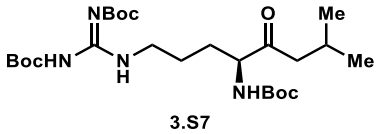
163.7, 156.4, 155.5, 153.5, 83.3, 79.9, 79.4, 57.4, 40.5, 37.7, 29.0, 28.5, 28.4, 28.2, 25.2, 19.2, 17.9; **IR** (thin film, cm^{-1}) 3331, 2976, 2932, 1718, 1640, 1617, 1415; **HRMS** (ESI) *m/z* calculated for $C_{24}H_{45}N_4O_7^+$ $[M+H]^+$ 501.3283, found 501.3285.



3.42d

(*S*)-1-(4-amino-6-methyl-5-oxoheptyl)guanidine (3.42d):

To a solution of tri-Boc protected arginine isopropyl ketone (**3.S6**, 6.1 mg, 0.012 mmol, 1.0 equiv) dissolved in DCM (0.4 mL) was added trifluoroacetic acid (TFA, 0.4 mL) at rt. The reaction was stirred for 5.3 h. After addition of dichloromethane (2.0 mL), the solution was evaporated under reduced pressure to afford 4.3 mg of the title compound as a bis-TFA salt with minor impurities (83% yield) as a colorless oil. R_f = 0.20 (RP_{18} TLC, 3:1 0.1% formic acid in H_2O/MeOH v/v); **^{1}H NMR** (600 MHz, CD_3OD) δ 4.37 (dd, *J* = 7.8, 4.0, 1H), 3.25 (dq, *J* = 12.2, 6.8, 2H), 2.96 (h, *J* = 6.8, 1H), 2.07 (ddt, *J* = 14.5, 12.1, 4.5, 1H), 1.85 (dddd, *J* = 14.6, 12.1, 7.8, 4.6, 1H), 1.77-1.69 (m, 1H), 1.66-1.58 (m, 1H), 1.18 (d, *J* = 7.0, 3H), 1.13 (d, *J* = 6.6, 3H); **^{13}C NMR** (150 MHz, CD_3OD) δ 210.6, 158.8, 58.0, 41.7, 38.3, 27.9, 25.4, 19.5, 17.6; **IR** (thin film, cm^{-1}) 3350, 3178, 2978, 1671; **HRMS** (ESI) *m/z* calculated for $C_9H_{21}N_4O^+$ $[M+H]^+$ 201.1710, found 201.1711.

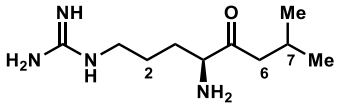


3.S7

Tri-Boc protected arginine isobutyl ketone (3.S7):

To a solution of tri-Boc-protected arginine Weinreb amide (**3.S1**, 50.0 mg, 0.0966 mmol, 1.0 equiv) dissolved in THF (1.9 mL) was added a 2 M solution of isobutyl magnesium bromide in

diethyl ether (725 μL, 1.45 mmol, 15.0 equiv) at 0 °C. The reaction was heated at 45 °C and stirred for 3.75 h. The reaction was poured into 15 mL of 0.1 M HCl. The acidified aqueous phase was extracted with EtOAc (3x). The organic layers were combined, washed with brine, dried over Na_2SO_4, filtered, and evaporated under reduced pressure to afford an oil. The crude residue was purified by flash column chromatography (4:1 hexanes/EtOAc v/v) to afford 10.0 mg of the title compound with minor impurities (20% yield) as a colorless oil. R_f = 0.56 (2:1 hexanes/EtOAc v/v); **^{1}H NMR** (600 MHz, $CDCl_3$) δ 11.47 (s, 1H), 8.33 (d, *J* = 5.9, 1H), 5.31 (d, *J* = 8.1, 1H), 4.29 (dd, *J* = 11.9, 5.6, 1H), 3.42 (dq, *J* = 12.2, 6.8, 5.6, 2H), 2.37 (t, *J* = 6.7, 2H), 2.17 (dp, *J* = 14.0, 6.9, 1H), 1.93-1.85 (m, 1H), 1.60 (dd, *J* = 11.6, 5.9, 1H), 1.58-1.50 (m, 2H), 1.49 (s, 9H), 1.49 (s, 9H), 1.43 (s, 9H), 0.92 (d, *J* = 6.8, 3H), 0.91 (d, *J* = 6.9, 3H); **^{13}C NMR** (150 MHz, $CDCl_3$) δ 208.9, 163.7, 156.4, 155.6, 153.5, 83.3, 79.7, 79.4, 59.2, 48.8, 40.4, 28.7, 28.5, 28.4, 28.2, 25.1, 24.5, 22.8, 22.7; **IR** (thin film, cm^{-1}) 3333, 2957, 2930, 2870, 1717, 1639, 1615, 1574, 1414; **HRMS** (ESI) *m/z* calculated for $C_{25}H_{47}N_4O_7^+$ $[M+H]^+$ 515.3439, found 515.3434.



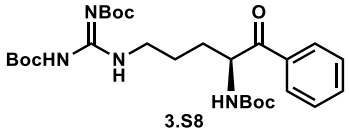
3.42e

(*S*)-1-(4-amino-7-methyl-5-oxooctyl)guanidine (3.42e):

To a solution of tri-Boc protected arginine isobutyl ketone (**3.S7**, 10.0 mg, 0.0194 mmol, 1.0 equiv) dissolved in DCM (0.5 mL) was added trifluoroacetic acid (TFA, 0.5 mL) at rt. The reaction was stirred for 3.3 h. After addition of dichloromethane (2.0 mL), the solution was evaporated under reduced pressure to afford 8.3 mg of the title compound as a bis-TFA salt with minor impurities (96% yield) as a colorless oil. R_f = 0.17 (RP_{18} TLC, 3:1 0.1% formic acid in H_2O/MeOH v/v); **^{1}H NMR** (600 MHz, CD_3OD) δ 4.15 (dd, *J* = 7.8, 4.1, 1H), 3.25 (td, *J* = 6.9, 4.4, 2H), 2.52 (d, *J* =

6.8, 2H), 2.19 (dp, *J* = 13.5, 6.7, 1H), 2.07 (ddt, *J* = 14.5, 12.1, 4.5, 1H), 1.85 (dddd, *J* = 14.6, 12.1, 7.8, 4.6, 1H), 1.77-1.67 (m, 1H), 1.67-1.58 (m, 1H), 0.97 (t, *J* = 6.4, 6H); **^{13}C NMR** (150 MHz, CD_3OD)* δ 206.3, 158.8, 59.8, 48.5, 41.7, 27.8, 25.0, 24.9, 22.7; **IR** (thin film, cm^{-1}) 3352, 3176, 2963, 2878, 1671, 1537, 1432; **HRMS** (ESI) *m/z* calculated for $C_{10}H_{23}N_4O^+$ $[M+H]^+$ 215.1866, found 215.1867.

*C2 and C7 were coalesced in ^{13}CNMR, as verified by HSQC. C6 was coalesced with the methanol signal in ^{13}CNMR and differentiated by HSQC.

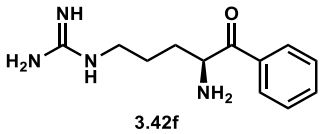


3.S8

Tri-Boc protected arginine phenyl ketone (3.S8):

To a solution of tri-Boc-protected arginine Weinreb amide (**3.S1**, 50.0 mg, 0.0966 mmol, 1.0 equiv) dissolved in THF (4.8 mL) was added a 3 M solution of phenyl magnesium bromide in diethyl ether (483 μL, 1.45 mmol, 15.0 equiv) at 0 °C. The reaction was allowed to warm to rt and stirred for 3.75 h. The reaction was poured into 15 mL of 0.1 M HCl. The acidified aqueous phase was extracted with EtOAc (3x). The organic layers were combined, washed with brine, dried over Na_2SO_4, filtered, and evaporated under reduced pressure to afford an oil. The crude residue was purified by flash column chromatography (4:1 hexanes/EtOAc v/v) to afford 39.9 mg of the title compound (77% yield) as a yellow oil. R_f = 0.54 (2:1 hexanes/EtOAc v/v); **^{1}H NMR** (600 MHz, $CDCl_3$) δ 11.44 (s, 1H), 8.26 (t, *J* = 5.2, 1H), 7.95 (d, *J* = 7.7, 2H), 7.58 (t, *J* = 7.4, 1H), 7.47 (t, *J* = 7.7, 2H), 5.54 (d, *J* = 8.2, 1H), 5.30 (td, *J* = 7.6, 4.1, 1H), 3.38 (q, *J* = 6.7, 2H), 1.94 (dd, *J* = 8.2, 4.7, 1H), 1.72-1.61 (m, 1H), 1.61-1.53 (m, 2H), 1.47 (s, 9H), 1.46 (s, 9H), 1.44 (s, 9H); **^{13}C NMR** (150 MHz, $CDCl_3$) δ 199.0, 163.7, 156.2, 155.6, 153.3, 134.6, 133.9, 129.0, 128.7, 83.2, 80.0,

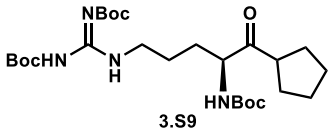
79.4, 54.8, 40.4, 30.8, 28.5, 28.4, 28.2, 24.9; **IR** (thin film, cm^{-1}) 3331, 2977, 2932, 1717, 1684, 1637, 1615, 1578, 1415; **HRMS** (ESI) *m/z* calculated for $C_{27}H_{43}N_4O_7^+$ $[M+H]^+$ 535.3126, found 535.3134.



3.42f

(*S*)-1-(4-amino-5-oxo-5-phenylpentyl)guanidine (3.42f):

To a solution of tri-Boc protected arginine phenyl ketone (**3.S8**, 39.9 mg, 0.0746 mmol, 1.0 equiv) dissolved in DCM (1.8 mL) was added trifluoroacetic acid (TFA, 1.8 mL) at rt. The reaction was stirred for 4.3 h. After addition of dichloromethane (2.0 mL), the solution was evaporated under reduced pressure to afford 34.5 mg of the title compound as a bis-TFA salt with minor impurities (quant) as a colorless oil. R_f = 0.19 (RP_{18} TLC, 3:1 0.1% formic acid in H_2O/MeOH v/v); **^{1}H NMR** (600 MHz, CD_3OD) δ 8.07-8.02 (m, 2H), 7.75-7.69 (m, 1H), 7.62-7.56 (m, 2H), 5.18 (dd, *J* = 7.7, 4.5, 1H), 3.17 (dp, *J* = 13.7, 6.9, 2H), 2.11-2.03 (m, 1H), 1.92 (dddd, *J* = 14.6, 12.1, 7.7, 4.7, 1H), 1.81-1.70 (m, 1H), 1.68-1.57 (m, 1H); **^{13}C NMR** (150 MHz, $CDCl_3$) δ 196.8, 158.7, 135.9, 134.7, 130.4, 129.9, 56.2, 41.6, 29.6, 25.2; **IR** (thin film, cm^{-1}) 3359, 3180, 2931, 1671, 1529, 1451, 1432; **HRMS** (ESI) *m/z* calculated for $C_{12}H_{19}N_4O^+$ $[M+H]^+$ 235.1553, found 235.1555.

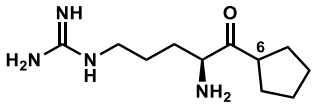


3.S9

Tri-Boc protected arginine cyclopentyl ketone (3.S9):

To a solution of tri-Boc-protected arginine Weinreb amide (**3.S1**, 50.0 mg, 0.0966 mmol, 1.0 equiv) dissolved in THF (1.9 mL) was added a 2 M solution of cyclopentyl magnesium bromide

in diethyl ether (725 μL, 1.45 mmol, 15.0 equiv) at 0 °C. The reaction was heated at 45 °C and stirred for 2.5 h. The reaction was poured into 15 mL of 0.1 M HCl. The acidified aqueous phase was extracted with EtOAc (3x). The organic layers were combined, washed with brine, dried over Na_2SO_4, filtered, and evaporated under reduced pressure to afford an oil. The crude residue was purified by flash column chromatography (4:1 hexanes/EtOAc v/v) to afford 13.0 mg of the title compound (26% yield) as a colorless oil. R_f = 0.58 (2:1 hexanes/EtOAc v/v); **^{1}H NMR** (600 MHz, $CDCl_3$) δ 11.47 (s, 1H), 8.32 (t, *J* = 5.5, 1H), 5.32 (d, *J* = 8.0, 1H), 4.43 (d, *J* = 7.3, 1H), 3.42 (td, *J* = 6.9, 5.3, 2H), 3.02 (p, *J* = 7.9, 1H), 1.95-1.88 (m, 2H), 1.84-1.71 (m, 3H), 1.71-1.60 (m, 4H), 1.60-1.55 (m, 3H), 1.49 (s, 9H), 1.48 (s, 9H), 1.43 (s, 9H); **^{13}C NMR** (150 MHz, $CDCl_3$) δ 212.1, 163.7, 156.4, 155.5, 153.5, 83.3, 79.8, 79.4, 58.6, 48.3, 40.5, 30.7, 28.9, 28.9, 28.5, 28.4, 28.2, 26.3, 26.3, 25.1; **IR** (thin film, cm^{-1}) 3332, 2976, 2869, 1717, 1639, 1616, 1574, 1415; **HRMS** (ESI) *m/z* calculated for $C_{26}H_{47}N_4O_7^+$ $[M+H]^+$ 527.3439, found 527.3443.



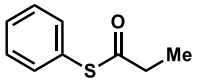
3.42g

(*S*)-1-(4-amino-5-cyclopentyl-5-oxopentyl)guanidine (3.42g):

To a solution of tri-Boc protected arginine cyclopentyl ketone (**3.S9**, 13.0 mg, 0.0247 mmol, 1.0 equiv) dissolved in DCM (0.7 mL) was added trifluoroacetic acid (TFA, 0.7 mL) at rt. The reaction was stirred for 4.3 h. After addition of dichloromethane (2.0 mL), the solution was evaporated under reduced pressure to afford 10.3 mg of the title compound as a bis-TFA salt with minor impurities (92% yield) as a colorless oil. R_f = 0.27 (RP_{18} TLC, 3:1 0.1% formic acid in H_2O/MeOH v/v); **^{1}H NMR** (600 MHz, CD_3OD) δ 4.29 (dd, *J* = 7.8, 4.1, 1H), 3.25 (td, *J* = 6.9, 5.5, 2H), 3.22-3.17 (m, 1H), 2.13-2.02 (m, 2H), 1.90-1.80 (m, 3H), 1.76-1.69 (m, 3H), 1.68-1.61 (m, 4H); **^{13}C**

NMR (150 MHz, CD_3OD)* δ 209.7, 158.8, 59.2, 48.2, 41.7, 31.9, 29.4, 27.7, 27.1, 27.1, 25.4; **IR** (thin film, cm^{-1}) 3367, 3180, 2958, 2874, 1671; **HRMS** (ESI) *m/z* calculated for $C_{11}H_{23}N_4O^+$ $[M+H]^+$ 227.1866, found 227.1868.

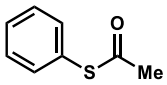
* The ^{13}C signal of C-6 coalesced with methanol, as elucidated by HSQC.



3.37

***S*-phenyl propanecarbothioate (3.37):**

To a suspension of anhydrous K_2CO_3 (502 mg, 3.63 mmol, 2.0 equiv) in dry EtOAc (9 mL) were added propionic anhydride (256 µL, 2.00 mmol, 1.1 equiv) and thiophenol (186 µL, 1.82 mmol, 1.0 equiv). The reaction mixture stirred under N_2 for 16 h. The reaction was quenched with 10 mL of water and extracted with EtOAc (3 x 10 mL). The combined organic layers were dried over Na_2SO_4, filtered and concentrated, affording 298 mg (99% yield) of the title compound as a clear colorless oil. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.46 – 7.36 (m, 5H), 2.69 (q, *J* = 7.5 Hz, 2H), 1.23 (t, *J* = 7.5 Hz, 3H). All spectra obtained were consistent with literature values.[75]

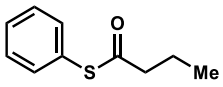


3.S10

***S*-phenyl ethanecarbothioate (3.S10):**

To a suspension of anhydrous K_2CO_3 (500 mg, 3.6 mmol, 2.0 equiv) in dry EtOAc (9 mL) were added acetic anhydride (189 µL, 2.0 mmol, 1.1 equiv) and thiophenol (186 µL, 1.8 mmol, 1.0 equiv). The reaction mixture stirred under N_2 for 16 h. The reaction was quenched with 10 mL of water and extracted with EtOAc (3 x 10 mL). The combined organic layers were dried over Na_2SO_4, filtered and concentrated, affording 270 mg (98% yield) of the title compound as a clear

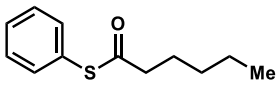
colorless oil. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.42 (d, *J* = 0.8 Hz, 5H), 2.43 (d, *J* = 0.8 Hz, 3H). All spectra obtained were consistent with literature values.[75]



3.S11

***S*-phenyl butanecarbothioate (3.S11):**

To a suspension of anhydrous K_2CO_3 (155 mg, 1.12 mmol, 2.0 equiv) in dry EtOAc (5 mL) were added butyric anhydride (100 µL, 0.62 mmol, 1.1 equiv) and thiophenol (58 µL, 0.56 mmol, 1.0 equiv). The reaction mixture stirred under N_2 for 16 h. The reaction was quenched with 5 mL of water and extracted with EtOAc (3 x 5 mL). The combined organic layers were dried over Na_2SO_4, filtered and concentrated, affording 88 mg (87% yield) of the title compound as a clear colorless oil. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.41 (apparent s, 5H), 2.64 (t, *J* = 7.4 Hz, 2H), 1.75 (h, *J* = 7.4 Hz, 2H), 1.00 (t, *J* = 7.4 Hz, 3H). All spectra were consistent with literature values.[76]

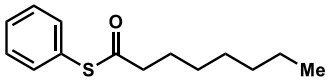


3.S12

***S*-phenyl hexanecarbothioate (3.S12):**

A solution of *N*-Ethyl-*N*′-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC, 118 mg, 0.76 mmol, 1.0 equiv) and hexanoic acid (100 µL, 0.80 mmol, 1.05 equiv) in DCM (10 mL) was cooled on ice for 15 min. Thiophenol (78 µL, 0.76 mmol, 1.0 equiv) and 4-(dimethylamino)pyridine (DMAP, 11 mg, 0.10 mmol, 0.13 equiv) were added to the cold solution, which was then allowed to warm to rt. The reaction mixture stirred under N_2 for 18 h and was

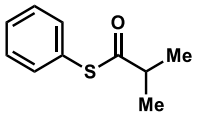
quenched with EtOAc (10 mL) and sat. aq. $NaHCO_3$ (20 mL). The quenched mixture was extracted with EtOAc (3 x 10 mL). The combined organic layers were washed with brine, dried over Na_2SO_4, filtered and concentrated under reduced pressure to an oil. The crude residue was purified by flash chromatography (2-20% gradient of DCM in hexanes) to afford 75 mg (47% yield) of the title compound as a colorless oil. R_f = 0.4 (20% DCM in hexanes). **^{1}H NMR** (600 MHz, CDCl3) δ 7.47 – 7.35 (m, 4H), 2.65 (t, *J* = 7.4 Hz, 2H), 1.72 (p, *J* = 7.4 Hz, 2H), 1.41 – 1.28 (m, 4H), 0.91 (t, *J* = 7.0 Hz, 3H). All spectra were consistent with literature values.[77]



3.S13

***S*-phenyl octanecarbothioate (3.S13):**

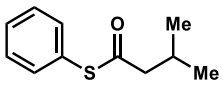
Thiophenol (513 µL, 5.0 mmol, 1.0 equiv) and triethylamine (697 µL, 5.0 mmol, 1.0 equiv) were dissolved in toluene. Octanoyl chloride (870 µL, 5.1 mmol, 1.02 equiv) was added dropwise over 40 min, forming a white precipitate. The reaction mixture stirred for 10 min after the addition of the acyl chloride and then was quenched with sat. aq. $NaHCO_3$ (15 mL). The organic layer was washed twice each with $NaHCO_3$ and brine (10 mL), then dried over Na_2SO_4, filtered. The filtrate was concentrated under reduced pressure affording 1.13 g (97% yield) of the title compound as a viscous clear oil. **^{1}H NMR** (600 MHz, $CDCl_3$) δ 7.41 (apparent s, 5H), 2.65 (t, *J* = 7.5 Hz, 2H), 1.71 (p, *J* = 7.4 Hz, 2H), 1.42 – 1.22 (m, 8H), 0.89 (t, *J* = 6.8 Hz, 3H). All spectra obtained were consistent with literature values.[78]



3.S14

***S*-phenyl 2-methylpropanecarbothioate (3.S14):**

To a suspension of anhydrous K_2CO_3 (155 mg, 1.12 mmol, 2.0 equiv) in dry EtOAc (5 mL) were added isobutyric anhydride (102 µL, 0.62 mmol, 1.1 equiv) and thiophenol (58 µL, 0.56 mmol, 1.0 equiv). The reaction mixture stirred under N_2 for 16 h. The reaction was quenched with 5 mL of water and extracted with EtOAc (3 x 5 mL). The combined organic layers were dried over Na_2SO_4, filtered and concentrated, affording 99 mg (99% yield) of the title compound as a clear colorless oil. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.41 (s, 5H), 2.87 (hept, *J* = 6.9 Hz, 1H), 1.27 (d, *J* = 6.9 Hz, 6H). All spectra obtained were consistent with literature values.[77]

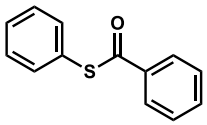


3.S15

***S*-phenyl 3-methylbutanecarbothioate (3.S15):**

Thiophenol (513 µL, 5.0 mmol, 1.0 equiv) and triethylamine (697 µL, 5.0 mmol, 1.0 equiv) were dissolved in toluene. Isovaleryl chloride (622 µL, 5.1 mmol, 1.02 equiv) was added dropwise over 40 min, forming a white precipitate. The reaction mixture stirred for 10 min after the addition of the acyl chloride and then was quenched with sat. aq. $NaHCO_3$ (15 mL). The organic layer was washed twice each with $NaHCO_3$ and brine (10 mL), then dried over Na_2SO_4, filtered. The filtrate was concentrated under reduced pressure affording 971 mg (99% yield) of the title compound as a viscous clear oil. **^{1}H NMR** (600 MHz, $CDCl_3$) δ 7.41 (d, *J* = 1.7 Hz, 5H), 2.54 (d, *J* = 7.1 Hz,

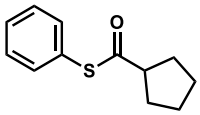
2H), 2.22 (hept, *J* = 13.5, 6.7 Hz, 1H), 1.01 (d, *J* = 6.7 Hz, 6H). All spectra obtained were consistent with literature values.[77]



3.S16

_S_-phenyl benzothioate (3.S16):

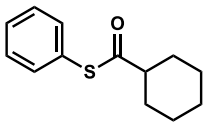
A solution of EDC (362 mg, 2.33 mmol, 1.0 equiv) and benzoic acid (300 mg, 2.45 mmol, 1.05 equiv) in DCM (25 mL) was cooled on ice for 15 min. Thiophenol (240 µL, 2.33 mmol, 1.0 equiv) and DMAP (36 mg, 0.29 mmol, 0.13 equiv) were added to the cold solution, which was then allowed to warm to rt. The reaction mixture stirred under N_2 for 18 h and was quenched with EtOAc (10 mL) and sat. aq. $NaHCO_3$ (20 mL). The quenched mixture was extracted with EtOAc (3 x 20 mL). The combined organic layers were washed with brine, dried over Na_2SO_4, filtered and concentrated under reduced pressure to an oil. The crude residue was purified by flash chromatography (2-20% gradient of DCM in hexanes) to afford 249 mg (50% yield) of the title compound as yellow/pink crystals. R_f = 0.4 (15% EtOAc in hexanes). **^{1}H NMR** (600 MHz, $CDCl_3$) δ 8.04 (dd, *J* = 8.1, 1.5 Hz, 2H), 7.61 (td, *J* = 7.4, 1.4 Hz, 1H), 7.55 – 7.43 (m, 7H). All spectra obtained were consistent with literature values.[79]



3.S17

***S*-phenyl cyclopentanecarbothioate (3.S17):**

Synthesized according to a variation of a procedure by Hansen et al.[57] To a solution of cyclopentane carboxylic acid (246 mg, 2.16 mmol, 1.0 equiv) and diphenyl disulfide (706 mg, 3.23 mmol, 1.5 equiv) dissolved in DCM (10 mL) was added tributylphosphine (808 µL, 3.23 mmol, 1.5 equiv) at 0 °C. The reaction warmed to rt and stirred for 3 h. The reaction was quenched by addition of sat. $NaHCO_3$ and extracted with DCM. The combined organic layers were dried over Na_2SO_4, filtered, and evaporated under reduced pressure to an oil. The crude residue was purified by flash column chromatography (4:1 hexanes/DCM v/v) to afford 357 mg of the title compound (80% yield) as a colorless oil. R_f = 0.2 (4:1 hexanes/DCM). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.45 – 7.36 (m, 5H), 3.10 (p, *J* = 7.9 Hz, 1H), 2.04 – 1.78 (m, 4H), 1.81 – 1.65 (m, 2H), 1.67 – 1.52 (m, 2H). All spectra obtained were consistent with literature values.[80]

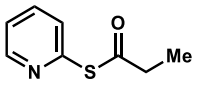


3.S18

***S*-phenyl cyclohexanecarbothioate (3.S18):**

Synthesized according to a variation of a procedure by Hansen et al.[57] To a solution of cyclohexanecarboxylic acid (247 mg, 2.16 mmol, 1.0 equiv) and diphenyl disulfide (706 mg, 3.23 mmol, 1.5 equiv) dissolved in DCM (10 mL) was added tributylphosphine (808 µL, 3.23 mmol, 1.5 equiv) at 0 °C. The reaction warmed to rt and stirred for 3 h. The reaction was quenched by addition of sat. $NaHCO_3$ and extracted with DCM. The combined organic layers were dried over

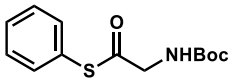
Na_2SO_4, filtered, and evaporated under reduced pressure to an oil. The crude residue was purified by flash column chromatography (5-15% v/v gradient of DCM in hexanes) to afford 348 mg of the title compound (73% yield) as a colorless oil. R_f = 0.2. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.40 (s, 5H), 2.61 (tt, *J* = 11.4, 3.5 Hz, 1H), 2.01 (ddt, *J* = 13.4, 3.3, 1.7 Hz, 2H), 1.82 (dt, *J* = 11.8, 3.1 Hz, 2H), 1.68 (dtd, *J* = 10.5, 3.3, 1.5 Hz, 1H), 1.60 – 1.46 (m, 2H), 1.40 – 1.16 (m, 3H). All spectra obtained were consistent with literature values.[79]



3.S19

***S*-(pyridin-2-yl) propanethioate (3.S19):**

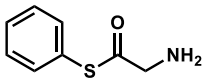
Synthesized according to a variation of a procedure by Hansen et al.[57] To a solution of propionic acid (200 mg, 2.7 mmol, 1.0 equiv) and 2,2'-dipyridyldisulfide (892 mg, 4.04 mmol, 1.5 equiv) dissolved in DCM (25 mL) was added tributylphosphine (1.01 mL, 4.04 mmol, 1.5 equiv) at 0 °C. The reaction warmed to rt and stirred for 3 h. The reaction was quenched by addition of sat. $NaHCO_3$ (20 mL) and extracted with DCM (3 x 15 mL). The combined organic layers were dried over Na_2SO_4, filtered, and evaporated under reduced pressure to an oil. The crude residue was purified by flash column chromatography (3-30% v/v gradient of EtOAc in hexanes) to afford 143 mg of the title compound (32% yield) as a yellow oil. R_f = 0.3. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.63 (ddd, J = 4.9, 2.0, 0.9 Hz, 1H), 7.75 (td, J = 7.7, 1.9 Hz, 1H), 7.62 (d, J = 7.9 Hz, 1H), 7.30 (ddd, J = 7.8, 4.9, 1.2 Hz, 1H), 2.73 (q, J = 7.5 Hz, 2H), 1.24 (t, J = 7.5 Hz, 4H). All spectra obtained were consistent with literature values.[81]



3.S20

***S*-phenyl 2-aminoethanethioate (Boc-Gly-SPh, 3.S20):**

A solution of EDC (253 mg, 1.63 mmol, 1.0 equiv) and Boc-glycine (300 mg, 1.71 mmol, 1.05 equiv) in DCM (20 mL) was cooled on ice for 15 min. Thiophenol (169 µL, 1.63 mmol, 1.0 equiv) and DMAP (12 mg, 0.11 mmol, 0.13 equiv) were added to the cold solution. The reaction warmed to rt, stirred for 16 hours under N_2. The reaction was quenched by the addition of sat. aq. $NaHCO_3$ (20 mL) and extracted with DCM (3 x 15 mL). The combined organic layers were dried over Na_2SO_4, filtered and concentrated under reduced pressure to an oil. The crude residue was purified by flash chromatography (10-20% v/v gradient of EtOAc in hexanes) to afford 307 mg (70% yield) of the title compound as an off-white solid. R_f = 0.3. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.42 (s, 5H), 5.11 (s, 1H), 4.15 (d, *J* = 6.1 Hz, 2H), 1.48 (s, 9H). All spectra obtained were consistent with literature values.[82]

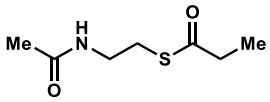


HCl

3.S21

***S*-phenyl 2-aminoethanethioate hydrochloride (Gly-SPh, 3.S21):**

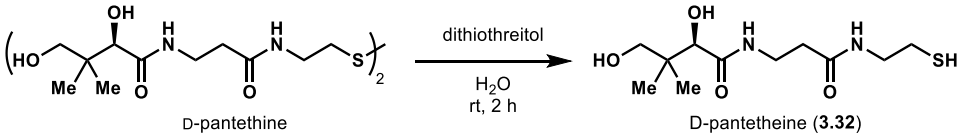
To a vial of thioester S3.20 (100 mg, 0.37 mmol) was added 2 mL of HCl in dioxane (4 M). The reaction mixture stirred for 2 h at rt. The liquids were removed under reduced pressure, affording 76 mg (quantitative yield) of the title compound as an off-white solid. **^{1}H NMR** (400 MHz, CD_3OD) δ 7.50 (s, 5H), 4.20 (s, 1H). All spectra obtained were consistent with literature values [82]

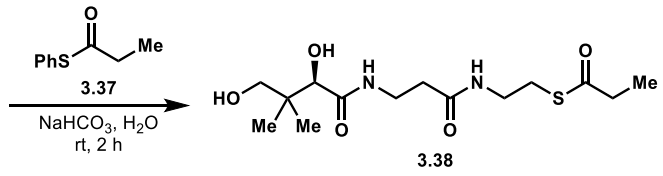


3.40

_S_-(2-acetamidoethyl) propanethioate (propionyl-SNAC, 3.40)

A solution of EDC (130 mg, 0.84 mmol, 1.0 equiv) and propionic acid (113 µL, 0.88 mmol, 1.05 equiv) in DCM (8 mL) was cooled on ice for 15 min. *N*-acetylcysteamine (90 µL, 0.84 mmol, 1.0 equiv, kindly donated by the group of Prof. David Sherman) and DMAP (12 mg, 0.11 mmol, 0.13 equiv) were added to the cold solution, which was then allowed to warm to rt. The reaction stirred for 12 hours and was then quenched by the addition of $CuSO_4$-impregnated silica (about 2 g). The slurry was filtered and extracted with a 1:1 v/v mixture of sat. aq. $NaHCO_3$ and brine (3 x 10 mL) The combined organic layers were dried over Na_2SO_4, filtered and concentrated under reduced pressure to a faintly blue oil. The crude residue was purified by flash chromatography (5% v/v MeOH in DCM) to afford 146 mg (99% yield) of the title compound as a yellow oil. R_f = 0.3 (5% MeOH in DCM). **^{1}H NMR** (400 MHz, $CDCl_3$) ^{1}H NMR (400 MHz, $CDCl_3$) δ 5.86 (s, 1H), 3.43 (q, *J* = 6.2 Hz, 2H), 3.02 (t, *J* = 6.4 Hz, 2H), 2.60 (q, *J* = 7.5 Hz, 2H), 1.96 (s, 3H), 1.18 (t, *J* = 7.5 Hz, 3H). All spectra obtained were consistent with literature values.[83]

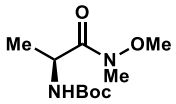




(*R*)-*S*-(2-(3-(2,4-dihydroxy-3,3-dimethylbutanamido)propanamido)ethyl) propanethioate (propionyl-pantetheine, 3.38):[53]

An approximately 1 M aqueous solution of D-pantethine (250 µL) was diluted in half with water. To the solution was added dithiothreitol (DTT, 60 mg, excess equiv), and the reaction mixture stirred under N_2 for 2 h. The liquid was removed under reduced pressure and azeotropic toluene drying. The remaining clear residue was purified by flash chromatography (2-10% v/v MeOH in DCM, R_f = 0.2), affording 112 mg of D-pantetheine (**3.32**) as a clear viscous oil that was stored overnight under N_2.

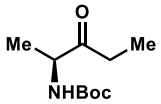
The purified pantetheine was redissolved in degassed and N_2-sparged water (10 mL). In a separate flask, propionyl-SPh (**3.37**, 87 mg, 1.3 equiv) was added to more degassed and sparged water (5 mL), forming a film on top. $NaCO_3$ (84 mg, 1.0 mmol, 2.5 equiv) was added in one portion to the aqueous mixture of **3.37**, followed by dropwise addition of the pantetheine solution over 10 min. The reaction mixture stirred for 2 h under sparging N_2 conditions. Enough NaCl to form a saturated solution was added to the reaction mixture, which was then extracted with THF (3 x 10 mL). The combined organic layers were dried over Na_2SO_4, filtered and concentrated under reduced pressure. The residue was purified over flash chromatography (2-10% v/v MeOH in DCM gradient), affording 106 mg (75% yield) of the title compound as a brown oil. R_f = 0.3 (10% MeOH in DCM). **^{1}H NMR** (600 MHz, $CDCl_3$) δ 7.49 (t, *J* = 6.2 Hz, 1H), 6.79 (t, *J* = 5.8 Hz, 1H), 4.62 (s, 1H), 4.05 (s, 1H), 3.97 (s, 1H), 3.52 (q, *J* = 6.5 Hz, 2H), 3.49 – 3.42 (m, 2H), 3.43 – 3.30 (m, 2H), 3.03 – 2.92 (m, 2H), 2.57 (q, *J* = 7.5 Hz, 2H), 2.41 (t, *J* = 6.1 Hz, 2H), 1.15 (t, *J* = 7.5 Hz, 3H), 0.96 (s, 3H), 0.89 (s, 3H); **^{13}C NMR** (150 MHz, $CDCl_3$) δ 200.9, 174.1, 171.9, 77.5, 70.89, 39.7, 39.4, 37.6, 35.8, 35.3, 28.3, 21.5, 20.6, 9.7; **HRMS** (ESI) *m/z* calculated for $C_{14}H_{26}N_2O_5SNa^+$ $[M+Na]^+$ 337.1455, found 357.1463.



3.S22

***tert*-butyl (*S*)-(1-(methoxy(methyl)amino)-1-oxopropan-2-yl)carbamate (3.S22)**

To a suspension of Boc-L-alanine (568 mg, 3.0 mmol, 1.0 equiv), *N*,*O*-dimethylhydroxyl-amine hydrochloride (265 mg, 3.3 mmol, 1.1 equiv), EDC (514 mg, 3.3 mmol, 1.1 equiv), HOBt (407 mg, 3.0 mmol, 1.0 equiv) in DCM (12 mL) was added triethylamine (838 µL, 6.0 mmol, 2.0 equiv). The yellow reaction mixture stirred for 18 h at rt under N_2. The reaction was quenched with sat. aq. $NaHCO_3$ (20 mL) and extracted with DCM (3 x 20 mL). The combined organic layers were washed once with more brine (20 mL), dried over Na_2SO_4, filtered and concentrated under reduced pressure. The crude residue (487 mg, 70% yield) was used without further purification. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 4.71 – 4.52 (m, 1H), 3.73 (s, 3H), 3.16 (s, 3H), 1.39 (s, 9H), 1.32 – 1.24 (m, 4H). All spectra obtained were consistent with literature values.[84]

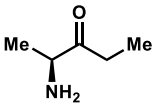


3.S23

***tert*-butyl (*S*)-(3-oxopentan-2-yl)carbamate (3.S23)**

To a solution of Boc-protected alanine Weinreb amide (**3.S22**, 400 mg, 1.72 mmol, 1.0 equiv) dissolved in THF (86 mL) was added a 1 M solution of $ZnCl_2$ in Et_2O (172 µL, 0.17 mmol, 0.1 equiv) and a 3 M solution of ethyl magnesium bromide in THF (8.6 mL, 25.8 mmol, 15.0 equiv) at 0 °C. The reaction was allowed to warm to rt and stirred for 4 h. The reaction was poured into 100 mL of 0.1 M HCl. The acidified aqueous phase was extracted with EtOAc (2 x 20 mL). The organic layers were combined, washed with brine, dried over Na_2SO_4, filtered, and evaporated under reduced pressure to afford an oil. The crude residue was purified by flash column chromatography (4:1 hexanes/EtOAc v/v) to afford 163 mg of the title compound (47% yield) as

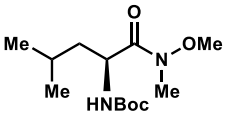
a yellow oil. **^{1}H NMR** (600 MHz, $CDCl_3$) δ 5.29 (d, *J* = 6.3 Hz, 1H), 4.27 (t, *J* = 7.2 Hz, 1H), 2.53 (dd, *J* = 17.9, 7.3 Hz, 1H), 2.44 (dd, *J* = 17.9, 7.3 Hz, 1H), 1.39 (s, 11H), 1.27 (d, *J* = 7.2 Hz, 4H), 1.04 (t, *J* = 7.3 Hz, 4H); **^{13}C NMR** (150 MHz, $CDCl_3$) δ 210.3, 155.3, 79.7, 54.9, 29.8, 28.4, 18.0, 7.67.



3.S24

(*S*)-2-aminopentan-3-one (3.S24)

To a solution of Boc-protected alanine ethyl ketone (**3.S23**, 101 mg, 0.50 mmol, 1.0 equiv) dissolved in DCM (1.7 mL) was added trifluoroacetic acid (TFA, 1.7 mL) at rt. The reaction was stirred for 2 h. The solvent was evaporated under reduced pressure to afford 46.7 mg of the title compound as a TFA salt (43% yield) as a colorless oil. **^{1}H NMR** (600 MHz, $CDCl_3$) δ 5.29 (d, *J* = 6.3 Hz, 1H), 4.27 (t, *J* = 7.2 Hz, 1H), 2.53 (dd, *J* = 17.9, 7.3 Hz, 1H), 2.44 (dd, *J* = 17.9, 7.3 Hz, 1H), 1.39 (s, 11H), 1.27 (d, *J* = 7.2 Hz, 4H), 1.04 (t, *J* = 7.3 Hz, 4H); **^{13}C NMR** (150 MHz, CD_3OD) δ 207.9, 55.6, 32.5, 15.74, 7.5; **HRMS** (ESI) *m/z* calculated for $C_5H_{12}NO^+$ $[M+H]^+$ 102.0913, found 102.0917.

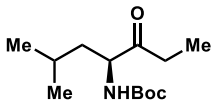


3.S25

***tert*-butyl (*S*)-(1-(methoxy(methyl)amino)-4-methyl-1-oxopentan-2-yl)carbamate (3.S25)**

To a suspension of Boc-L-leucine hydrate (250 mg, 1.0 mmol, 1.0 equiv), *N*,*O*-dimethylhydroxyl-amine hydrochloride (88 mg, 1.1 mmol, 1.1 equiv), EDC (171 mg, 1.1 mmol, 1.1 equiv), hydroxybenzotriazole (HOBt, 135 mg, 1.0 mmol, 1.0 equiv) in DCM (4 mL) was added diisopropylethyl amine (280 µL, 2.0 mmol, 2.0 equiv). The yellow reaction mixture stirred for 18

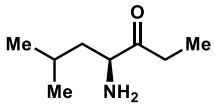
h at rt under N_2. The reaction was quenched with sat. aq. NH_4Cl (10 mL) and extracted with DCM (3 x 10 mL). The combined organic layers were washed once with more NH_4Cl (10 mL), dried over Na_2SO_4, filtered and concentrated under reduced pressure. The crude residue (150 mg, 54% yield) was used without further purification. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 5.06 (d, *J* = 9.6 Hz, 1H), 4.73 – 4.63 (m, 1H), 3.75 (s, 3H), 3.16 (s, 3H), 1.68 (dp, *J* = 13.1, 6.5 Hz, 1H), 1.39 (s, 9H), 0.91 (dd, *J* = 14.4, 6.6 Hz, 6H). All spectra obtained were consistent with literature values.[85]



3.S26

***tert*-butyl (*S*)-(2-methyl-5-oxoheptan-4-yl)carbamate (3.S26)**

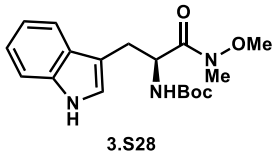
To a solution of tri-Boc-protected leucine Weinreb amide (**3.S25**, 120 mg, 0.44 mmol, 1.0 equiv) dissolved in THF (22 mL) was added a 1 M solution of $ZnCl_2$ (44 µL, 0.04 mmol, 0.1 equiv) and a 3 M solution of ethyl magnesium bromide in diethyl ether (2.2 mL, 6.6 mmol, 15.0 equiv) at 0 °C. The reaction was allowed to warm to rt and stirred for 4 h. The reaction was poured into 20 mL of 0.1 M HCl. The acidified aqueous phase was extracted with EtOAc (3x). The organic layers were combined, washed with brine, dried over Na_2SO_4, filtered, and evaporated under reduced pressure to afford an oil. The crude residue was purified by flash column chromatography (10% v/v hexanes in EtOAc) to afford 83.0 mg of the title compound (69% yield) as a yellow oil. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 5.02 (d, *J* = 8.3 Hz, 1H), 4.31 (td, *J* = 9.1, 3.9 Hz, 1H), 2.63 – 2.37 (m, 2H), 1.74 – 1.61 (m, 1H), 1.41 (s, 9H), 1.09 – 1.02 (m, 3H), 0.93 (dd, *J* = 17.7, 6.6 Hz, 6H).



3.S27

(*S*)-4-amino-6-methylheptan-3-one (3.S27)

To a solution of Boc-protected leucine ethyl ketone (**3.S26**, 70.0 mg, 0.29 mmol, 1.0 equiv) dissolved in DCM (2.0 mL) was added trifluoroacetic acid (TFA, 5.0 mL) at rt. The reaction was stirred for 2 h. The solvent was evaporated under reduced pressure to afford 26.9 mg of the title compound as a TFA salt (37% yield) as a colorless oil. **^{1}H NMR** (400 MHz, CD_3OD) δ 4.13 (dd, *J* = 9.9, 3.2 Hz, 1H), 2.71 (dq, *J* = 18.4, 7.2 Hz, 1H), 2.56 (dq, *J* = 18.4, 7.2 Hz, 1H), 1.77 (qq, *J* = 6.2, 3.5 Hz, 2H), 1.59 (t, *J* = 9.8 Hz, 1H), 1.10 (t, *J* = 7.2 Hz, 3H), 1.03 (dd, *J* = 10.2, 6.1 Hz, 6H); **^{13}C NMR** (151 MHz, CD_3OD) δ 207.9, 58.3, 39.9, 33.2, 25.7, 23.4, 21.4, 7.6; **HRMS** (ESI) *m/z* calculated for $C_8H_{18}NO^+$ $[M+H]^+$ 144.1383, found 144.1387.

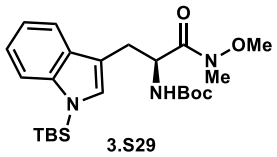


3.S28

***tert*-butyl (*S*)-(3-(1H-indol-3-yl)-1-(methoxy(methyl)amino)-1-oxopropan-2-yl)carbamate (3.S28)**

To a suspension of Boc-L-tryptophan (250 mg, 0.82 mmol, 1.0 equiv), *N*,*O*-dimethylhydroxyl-amine hydrochloride (132 mg, 1.35 mmol, 1.65 equiv), EDC (189 mg, 0.99 mmol, 1.2 equiv), HOBt (134 mg, 0.99 mmol, 1.2 equiv) and DMAP (13 mg, 0.10 mmol, 0.13 equiv) in DCM (10 mL) was added diisopropylethyl amine (859 µL, 4.9 mmol, 6.0 equiv). The yellow reaction mixture stirred for 18 h at rt under N_2. The reaction was quenched with sat. aq. NH_4Cl (10 mL) and extracted with DCM (3 x 10 mL). The combined organic layers were washed once with more NH_4Cl (about 20 mL), dried over Na_2SO_4, filtered and concentrated under reduced pressure. The crude residue was purified by flash chromatography (33-45% v/v gradient of EtOAc in hexanes)

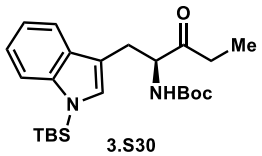
to afford 150 mg (53% yield) of a clear oil. R_f = 0.2 (40% EtOAc in hexanes). **^{1}H NMR** (600 MHz, cdcl$_3$) δ 8.01 (s, 1H), 7.60 (d, *J* = 7.9 Hz, 1H), 7.34 (d, *J* = 8.1 Hz, 1H), 7.18 (t, *J* = 7.5 Hz, 1H), 7.11 (t, *J* = 7.5 Hz, 1H), 7.06 (d, *J* = 2.3 Hz, 1H), 5.26 – 5.21 (m, 1H), 5.04 – 4.98 (m, 1H), 3.65 (s, 3H), 3.26 – 3.19 (m, 1H), 3.15 (s, 3H), 1.40 (s, 9H). All spectra obtained were consistent with literature values.[86]



3.S29

***tert*-butyl (*S*)-(3-(1-(tert-butyldimethylsilyl)-1H-indol-3-yl)-1-(methoxy(methyl)amino)-1-oxopropan-2-yl)carbamate (3.S29)**

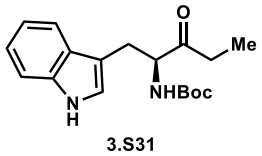
A solution of Boc-protected tryptophan Weinreb amide (**3.S28**, 150 mg, 0.43 mmol, 1.0 equiv) in THF (3.5 mL) was cooled to -78 °C and stirred for 1 h with lithium hexamethyldisilazide (1.38 mL of a 1 M solution in THF, 1.38 mmol, 3.2 equiv) under N_2. *Tert*-butyldimethylsilyl chloride (78 mg, 0.52 mmol, 1.2 equiv) was added in one portion and the reaction mixture stirred for a further 2 h at -78 °C. After warming to rt, the reaction was quenched by the addition of sat. aq. NH_4Cl (5 mL). The organic solvent was removed under reduced pressure, and the aqueous solution was extracted with EtOAc (3 x 10 mL). The combined organic layers were dried over Na_2SO_4, filtered and concentrated under reduced pressure. The crude residue was purified by flash chromatograph (30-40% v/v gradient of EtOAc in hexanes) to afford 157 mg (79% yield) of the title compound as a white foam. R_f = 0.3 (30% EtOAc in hexanes). **^{1}H NMR** (600 MHz, $CDCl_3$) δ 7.57 (d, *J* = 7.5 Hz, 1H), 7.45 (d, *J* = 7.9 Hz, 1H), 7.12 (p, *J* = 7.0 Hz, 2H), 7.00 (s, 1H), 5.21 (d, *J* = 8.7 Hz, 1H), 5.01 (d, *J* = 8.2 Hz, 1H), 3.57 (s, 3H), 3.22 – 3.17 (m, 1H), 3.12 (s, 3H), 1.40 (s, 6H), 0.91 (s, 9H), 0.57 (s, 6H). All spectra obtained were consistent with literature values.[86]



3.S30

***tert*-butyl (*S*)-(1-(1-(tert-butyldimethylsilyl)-1H-indol-3-yl)-3-oxopentan-2-yl)carbamate (3.S30)**

To a solution of Boc- and TBS-protected amide (3.S29, 157 mg, 0.34 mmol, 1.0 equiv) in THF (3 mL) cooled to 0 °C, was added ethyl magnesium bromide (340 µL of a 3 M solution in dimethyl ether, 1.02 mmol, 3.0 equiv) over 5 min. The reaction mixture stirred at 0 °C for 30 min, then warmed to rt over 1.5 h. The reaction was quenched by the addition of sat. aq. NH_4Cl (about 5 mL) and extracted with EtOAc (3 x 5 mL). The combined organic layers were dried over Na_2SO_4, filtered and concentrated under reduced pressure. The crude residue was purified by flash chromatography (10-20% v/v gradient of EtOAc in hexanes) to afford 114 mg (78% yield) of the title compound as an off-white solid. R_f = 0.5 (15% EtOAc in hexanes). **^{1}H NMR** (600 MHz, $CDCl_3$) δ 7.60 (dd, *J* = 7.2, 1.5 Hz, 1H), 7.49 (d, *J* = 8.3 Hz, 1H), 7.21 – 7.09 (m, 2H), 6.95 (d, *J* = 2.9 Hz, 1H), 5.26 (d, *J* = 7.8 Hz, 1H), 4.65 (q, *J* = 6.8 Hz, 1H), 3.19 (dd, *J* = 6.3, 2.7 Hz, 2H), 2.35 (pd, *J* = 9.9, 8.7, 3.8 Hz, 2H), 1.45 (s, 5H), 0.98 – 0.91 (m, 13H), 0.59 (t, *J* = 2.5 Hz, 7H); **^{13}C NMR** (150 MHz, $CDCl_3$) δ 210.7, 155.4, 141.4, 130.9, 129.4, 121.2, 119.9, 118.9, 114.0, 112.5, 79.7, 59.1, 34.3, 28.4, 28.1, 26.4, 19.5, 7.3, -3.9.

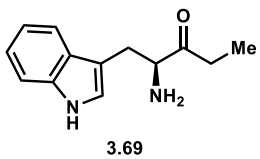


3.S31

***tert*-butyl (*S*)-(1-(1H-indol-3-yl)-3-oxopentan-2-yl)carbamate (3.S31)**

To a solution of Boc- and TBS-protected ketone (**3.S30**, 40 mg, 0.093 mmol, 1.0 equiv) in THF (1.8 mL) was added TBAF (464 µL of a 1 M solution, 0.46 mmol, 5.0 equiv). The reaction stirred

for 45 min and was quenched by the addition of sat. aq. NH_4Cl (2 mL) and then extracted with EtOAc (3 x 5 mL). The combined organic layers were dried over Na_2SO_4, filtered and concentrated under reduced pressure. The crude residue was purified by flash chromatography (25-35% v/v gradient of EtOAc in hexanes) to afford 23.7 mg (80% yield) of the title compound as a clear oil. R_f = 0.2 (20% EtOAc in hexanes). **^{1}H NMR** (600 MHz, $CDCl_3$) δ 8.24 (d, *J* = 16.4 Hz, 1H), 7.59 (d, *J* = 8.0 Hz, 1H), 7.35 (d, *J* = 8.1 Hz, 1H), 7.20 (t, *J* = 7.6 Hz, 1H), 7.13 (t, *J* = 7.5 Hz, 1H), 6.96 (s, 1H), 5.25 (d, *J* = 7.5 Hz, 1H), 4.64 (q, *J* = 6.8 Hz, 1H), 3.21 (t, *J* = 5.4 Hz, 1H), 2.38 (q, *J* = 7.3 Hz, 2H), 1.43 (s, 9H), 0.96 (t, *J* = 7.2 Hz, 3H); **^{13}C NMR** (150 MHz, $CDCl_3$) δ 210.76, 155.50, 136.25, 127.62, 122.72, 122.39, 119.82, 119.80, 118.94, 111.33, 110.50, 79.91, 77.37, 77.16, 76.95, 59.47, 34.11, 28.47, 27.96, 7.49.

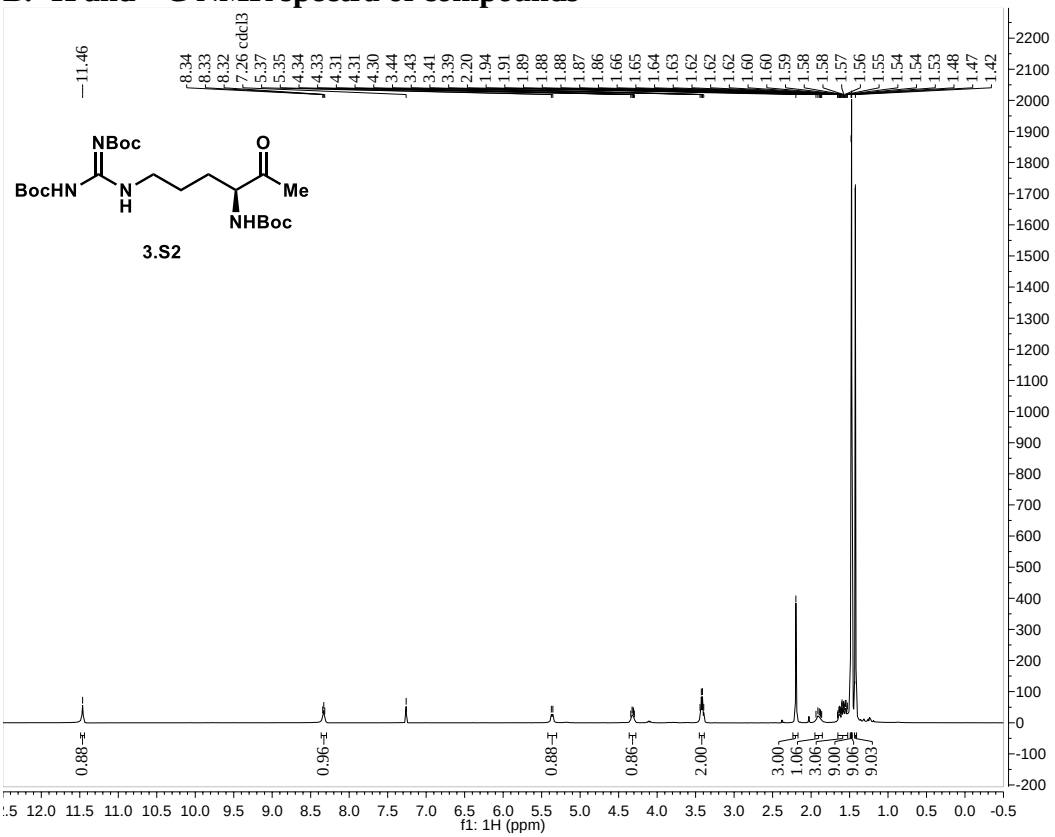


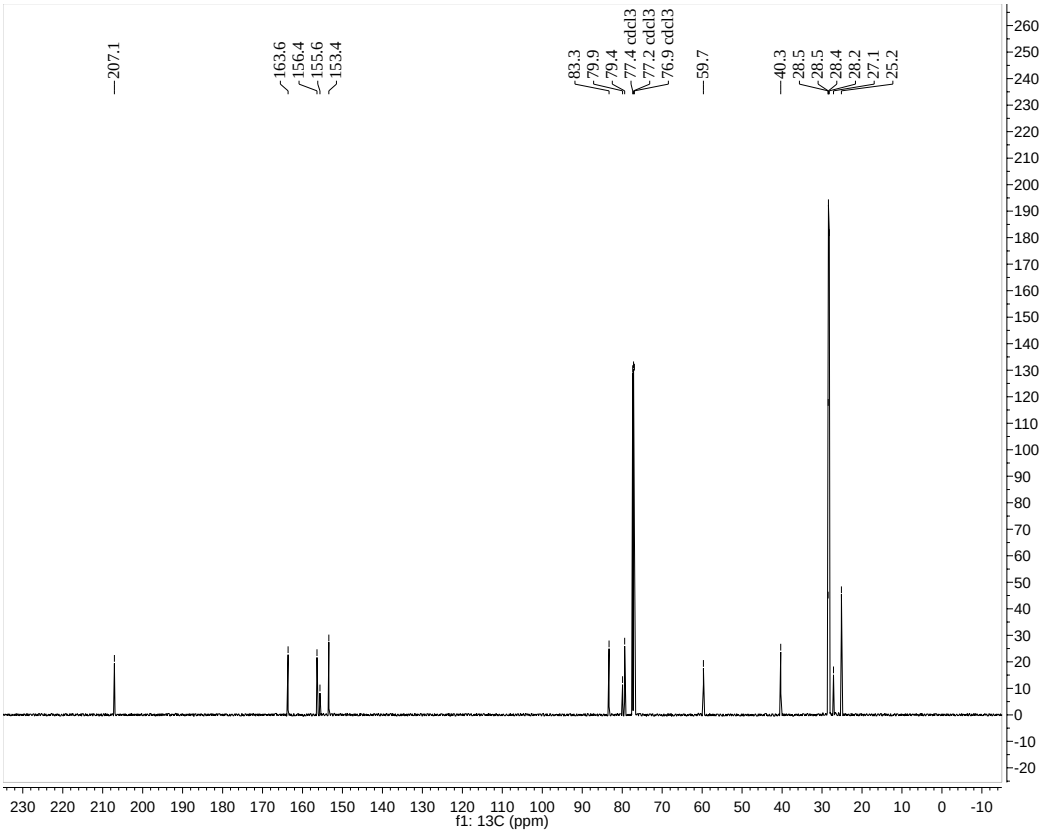
3.69

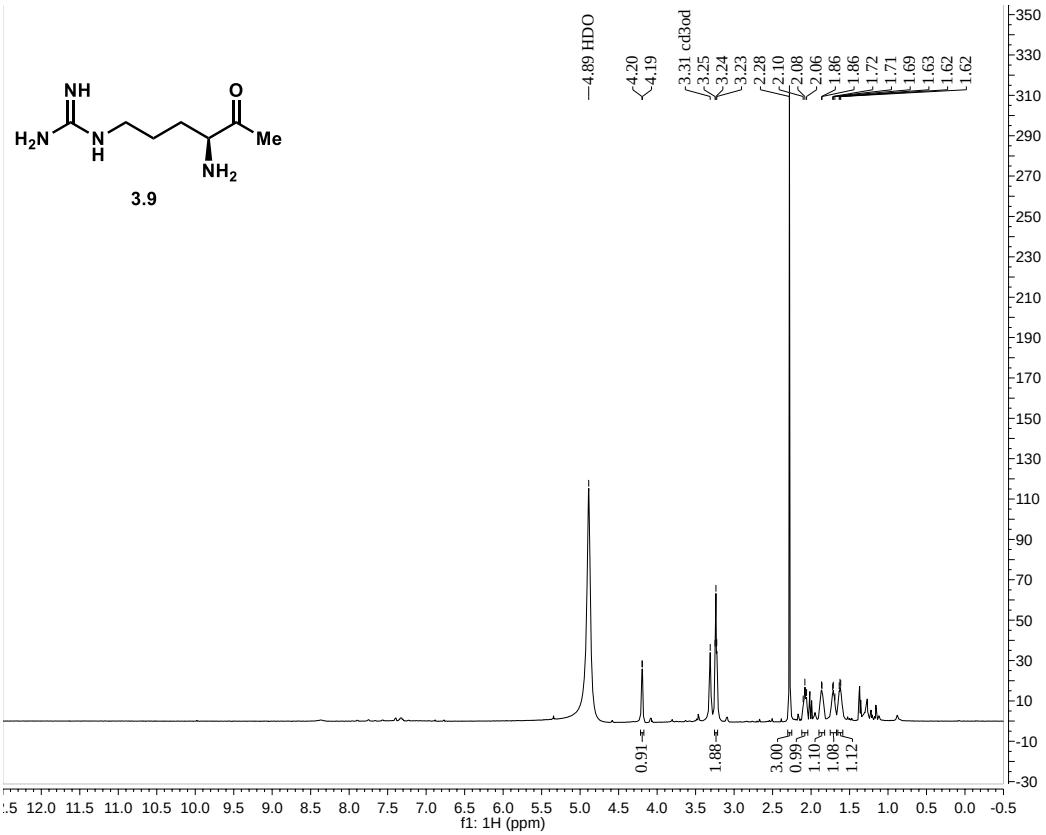
(*S*)-2-amino-1-(1H-indol-3-yl)pentan-3-one (3.69)

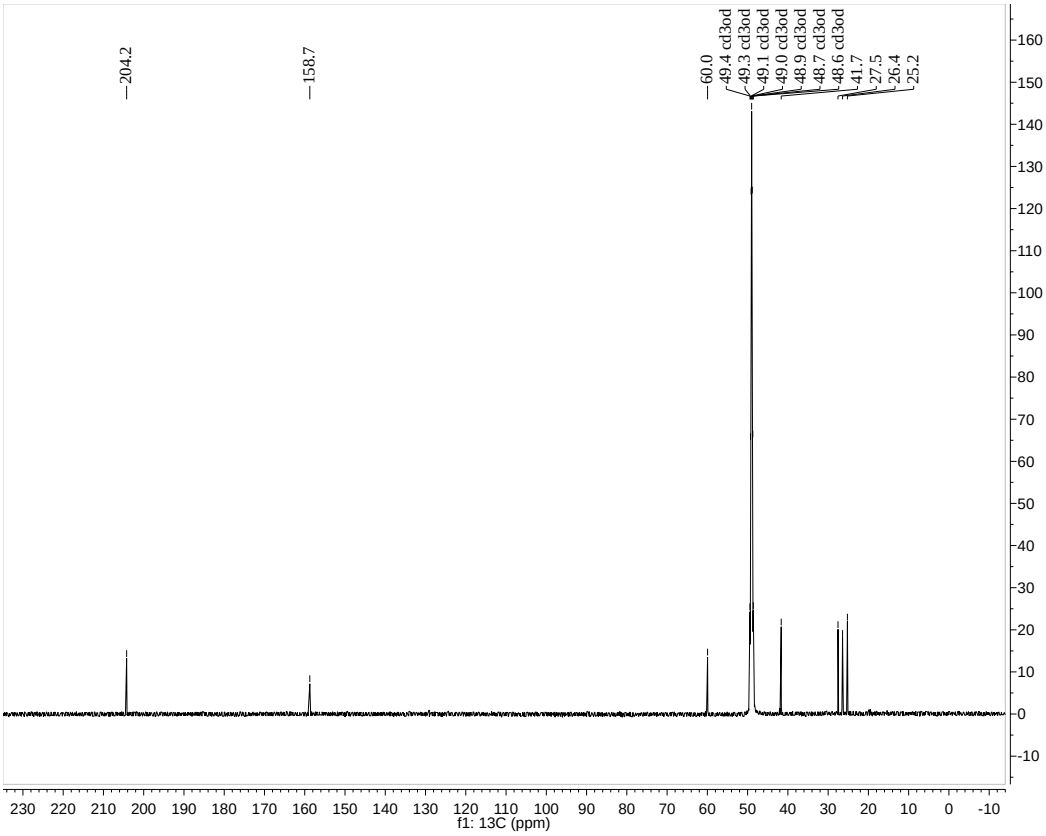
To a solution of Boc-protected tryptophan ethyl ketone (**3.S31**, 16.3 mg, 0.52 mmol, 1.0 equiv) dissolved in DCM (1.7 mL) was added trifluoroacetic acid (TFA, 1.7 mL) at rt. The reaction was stirred for 1 h. The solvent was evaporated under reduced pressure to afford 17.4 mg of the title compound as a TFA salt (quantitative yield) as a colorless oil. **^{1}H NMR** (600 MHz, CD_3OD) δ 7.60 (d, *J* = 7.9 Hz, 1H), 7.40 (d, *J* = 8.2 Hz, 1H), 7.19 (s, 1H), 7.15 (d, *J* = 8.1 Hz, 1H), 7.09 (t, *J* = 7.5 Hz, 1H), 4.43 (dd, *J* = 8.1, 6.5 Hz, 1H), 3.44 (dd, *J* = 14.9, 6.5 Hz, 1H), 3.21 (dd, *J* = 14.9, 8.1 Hz, 1H), 2.52 (q, *J* = 7.2 Hz, 2H), 1.00 (t, *J* = 7.2 Hz, 3H); **^{13}C NMR** (151 MHz, CD_3OD) δ 207.9, 138.4, 128.3, 125.3, 123.1, 120.4, 118.9, 112.7, 107.9, 59.9, 34.5, 27.6, 7.4; **HRMS** (ESI) *m/z* calculated for $C_{13}H_{17}N_2O^+$ $[M+H]^+$ 217.1335, found 217.1339.

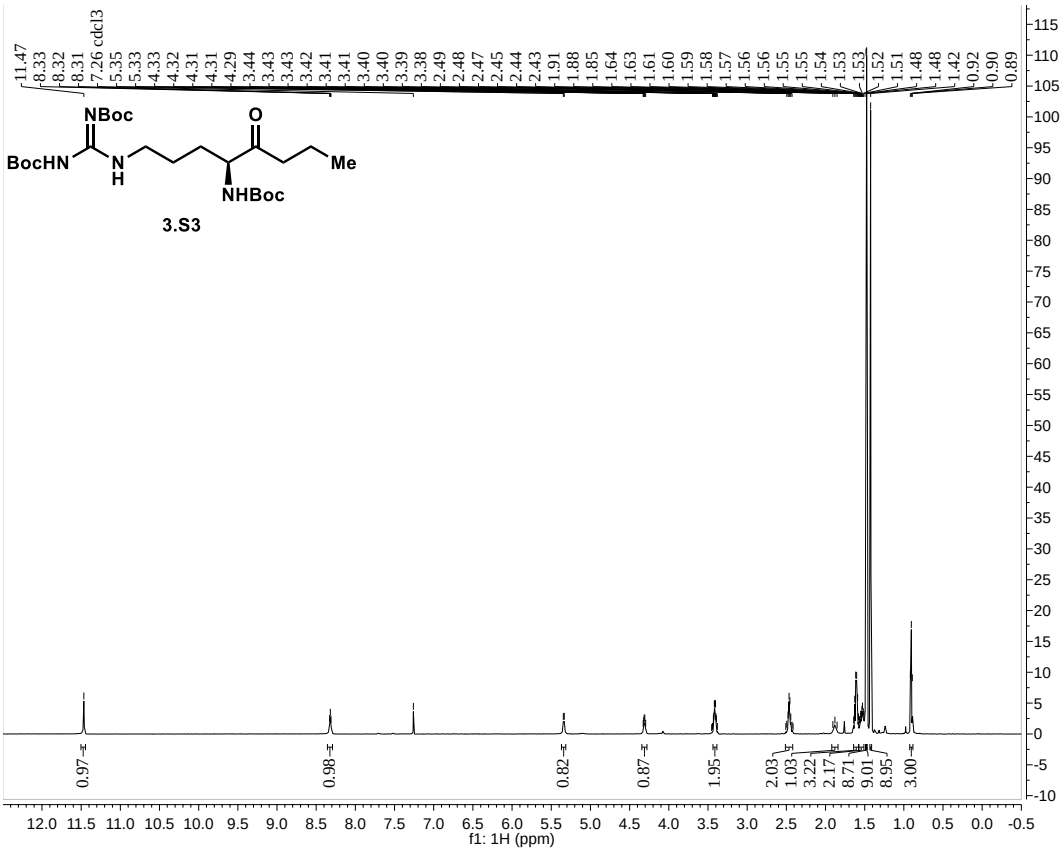
B. ^{1}H and ^{13}C NMR spectra of compounds

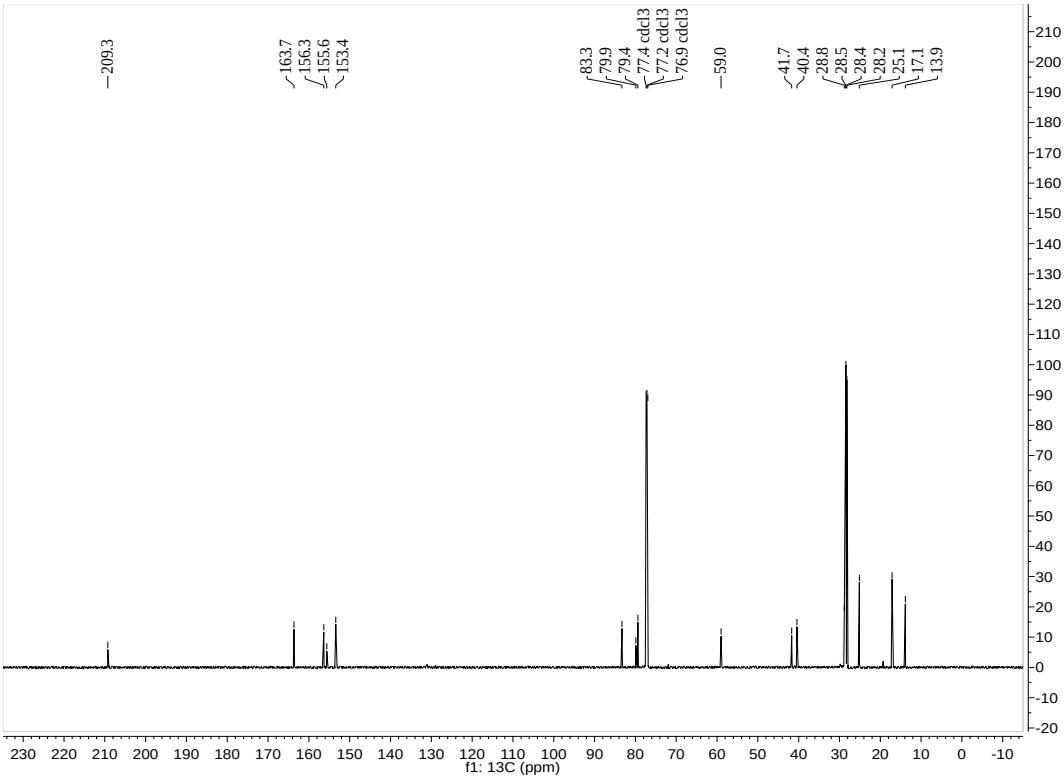


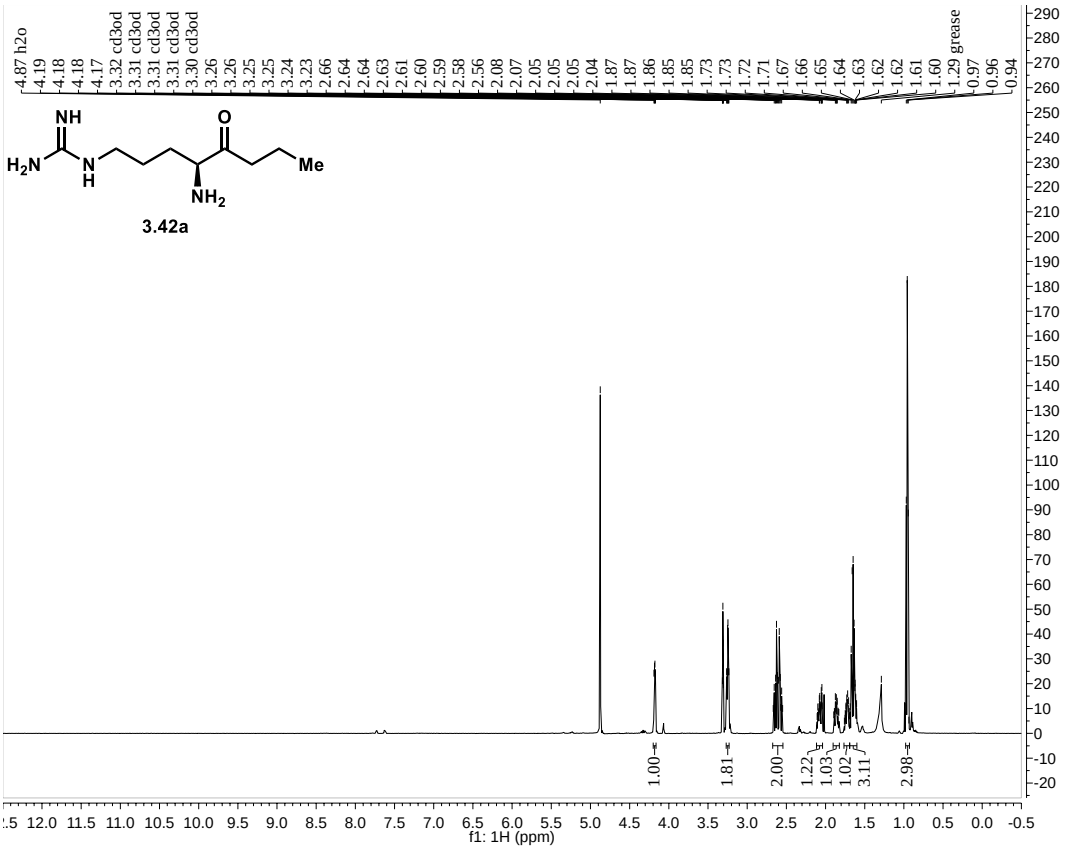


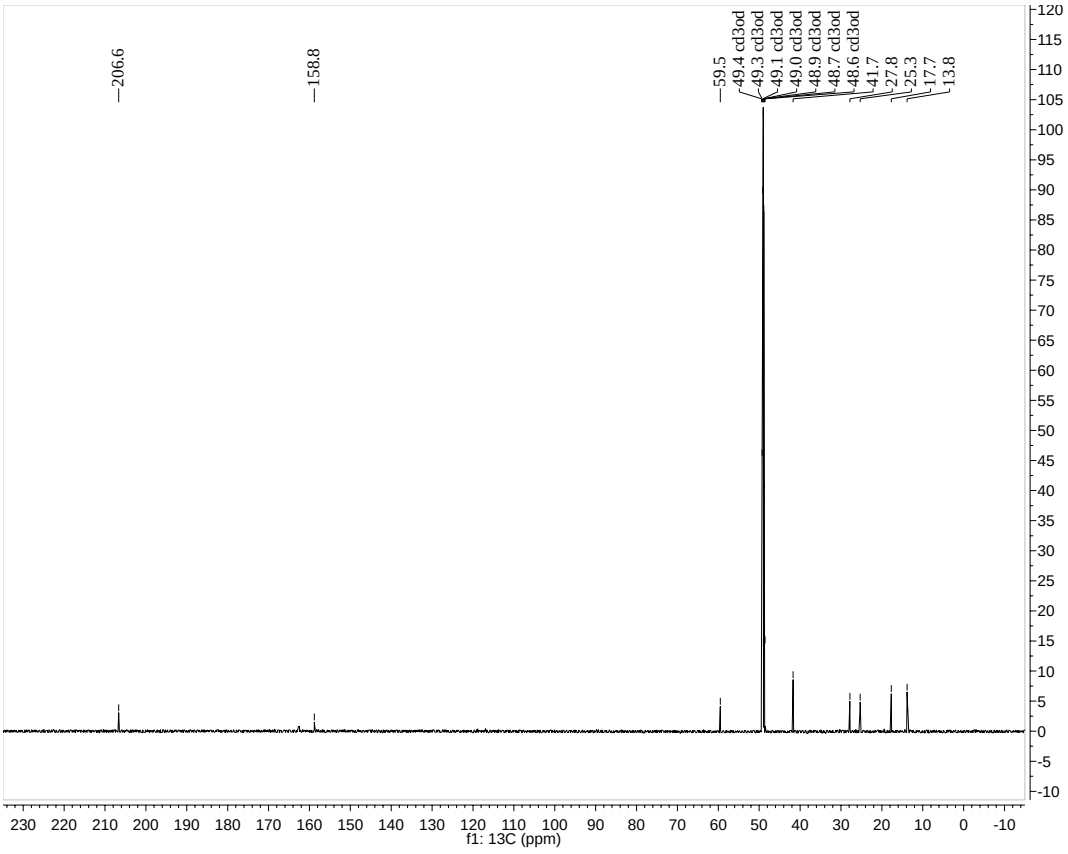

NH
O
H2N
N
H
Me
NH2
3.9
4.89 HDO
4.20
4.19
3.31 cd3od
3.25
3.24
3.23
2.28
2.10
2.08
2.06
1.86
1.86
1.72
1.71
1.69
1.63
1.62
1.62
0.91
1.88
3.00
0.99
1.10
1.08
1.12
f1: 1H (ppm)

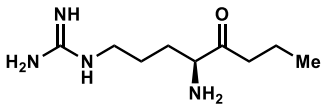

204.2
158.7
60.0
49.4 cd3od
49.3 cd3od
49.1 cd3od
49.0 cd3od
48.9 cd3od
48.7 cd3od
48.6 cd3od
41.7
27.5
26.4
25.2
f1: 13C (ppm)


NBoc
O
BocHN
N
H
Me
NHBoc
3.S3
f1: 1H (ppm)


f1: 13C (ppm)

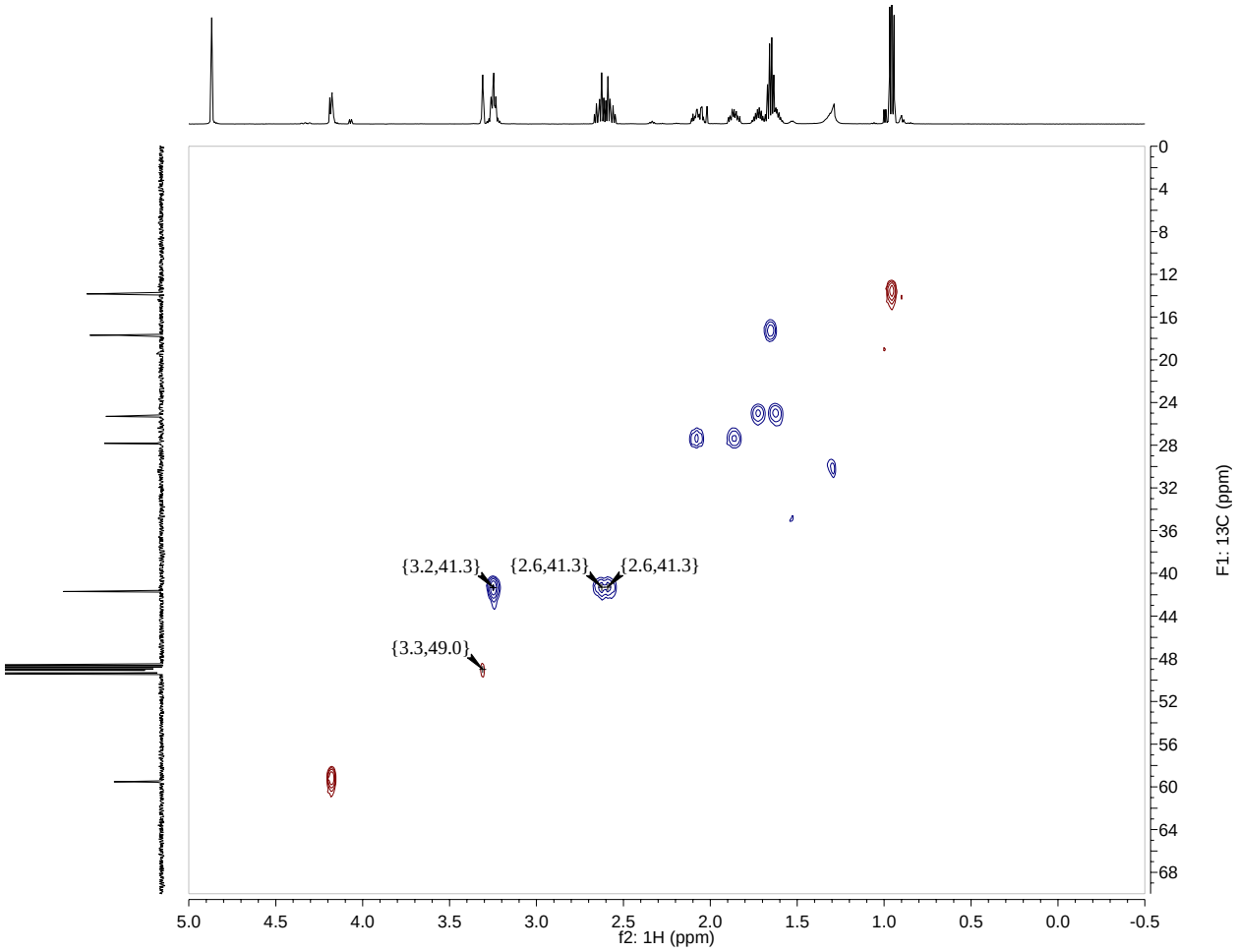

4.87 h2o
4.19
4.18
4.18
4.17
3.32 cd3od
3.31 cd3od
3.31 cd3od
3.31 cd3od
3.30 cd3od
3.26
3.26
3.25
3.25
3.24
3.23
2.66
2.64
2.64
2.63
2.61
2.60
2.59
2.58
2.56
2.08
2.07
2.05
2.05
2.04
1.87
1.87
1.86
1.85
1.85
1.73
1.73
1.72
1.71
1.67
1.66
1.65
1.64
1.63
1.62
1.62
1.61
1.60
1.29 grease
0.97
0.96
0.94
NH
O
H2N
N
H
Me
NH2
3.42a
1.00
1.81
2.00
1.22
1.03
1.02
3.11
2.98
f1: 1H (ppm)

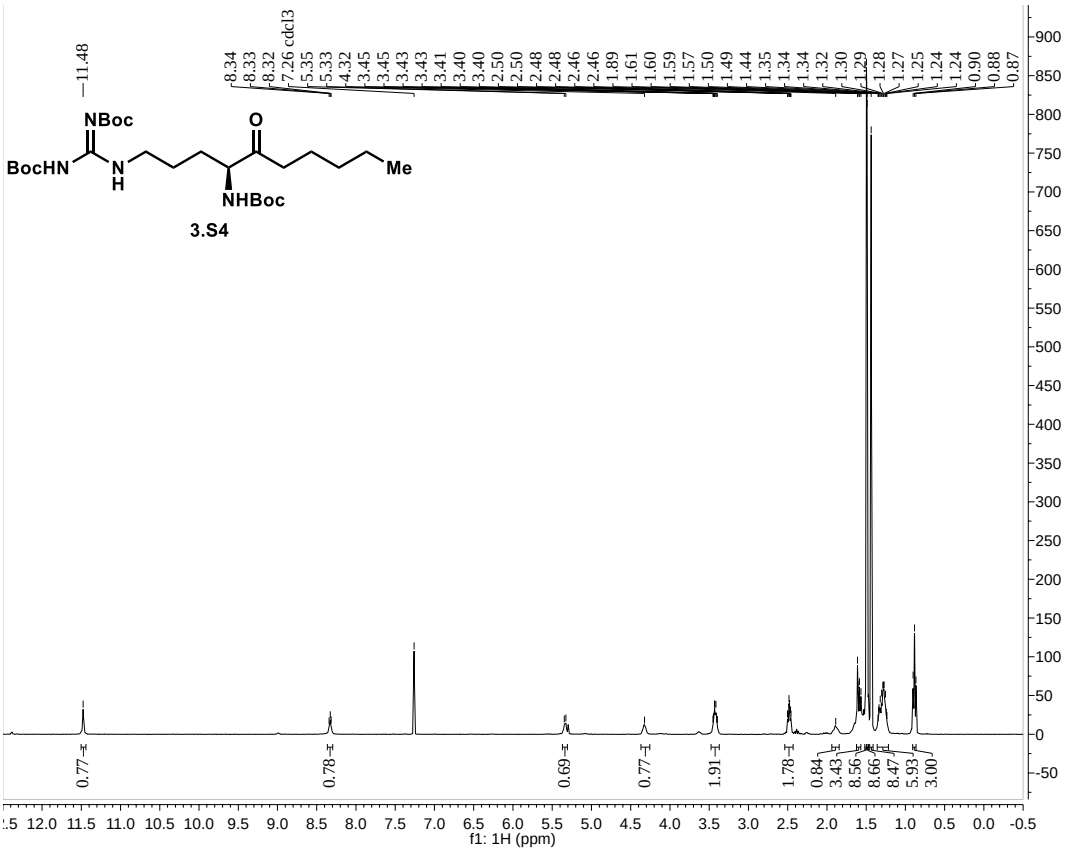

206.6
158.8
59.5
49.4 cd3od
49.3 cd3od
49.1 cd3od
49.0 cd3od
48.9 cd3od
48.7 cd3od
48.6 cd3od
41.7
27.8
25.3
17.7
13.8
f1: 13C (ppm)

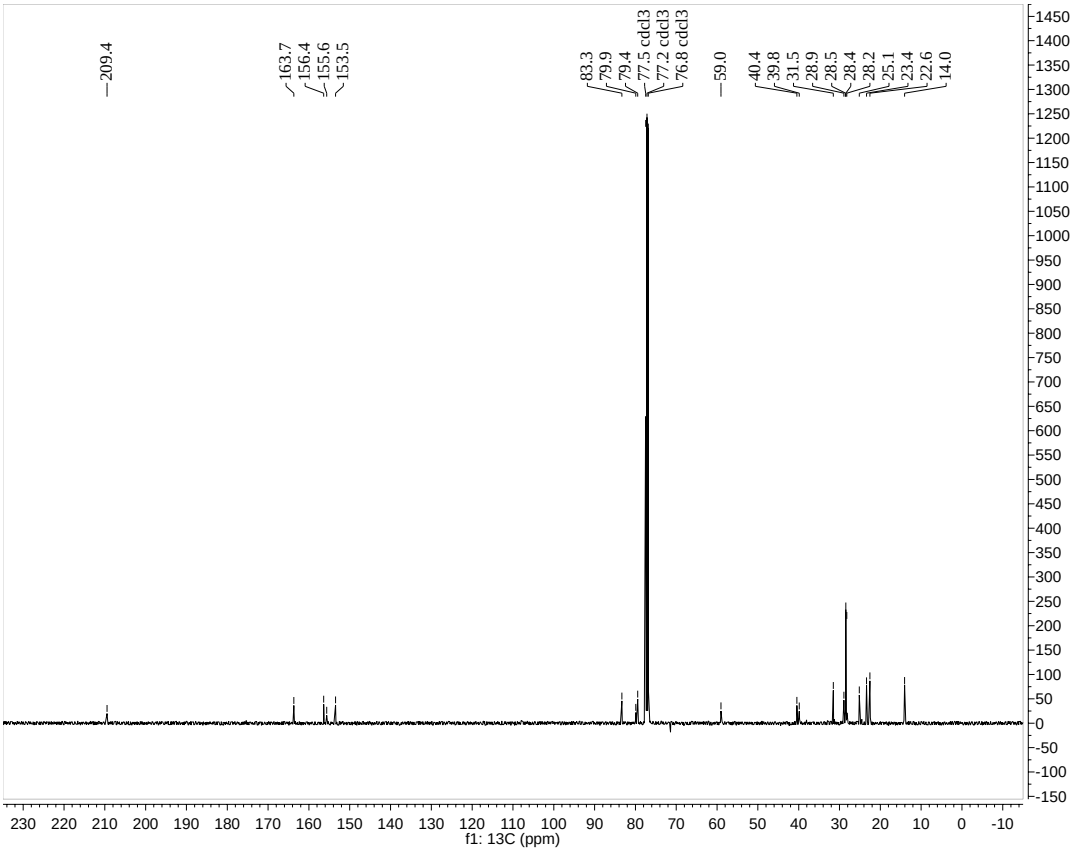


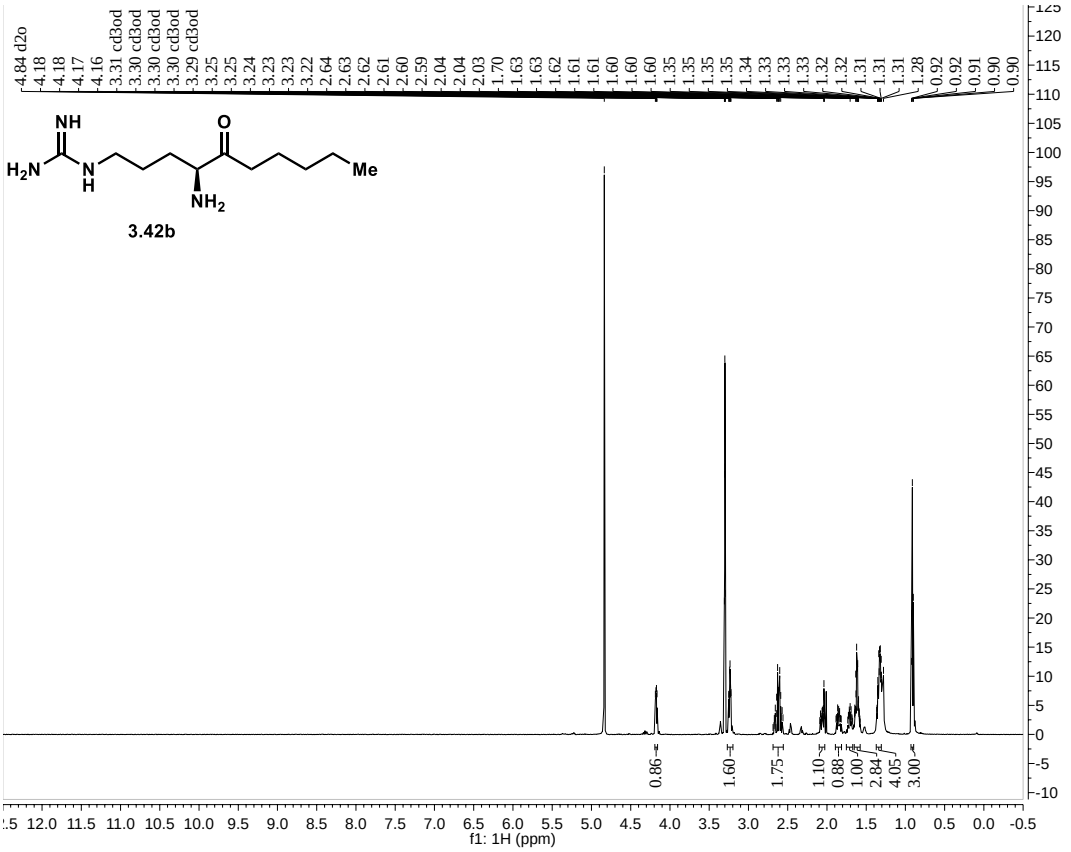
3.42a

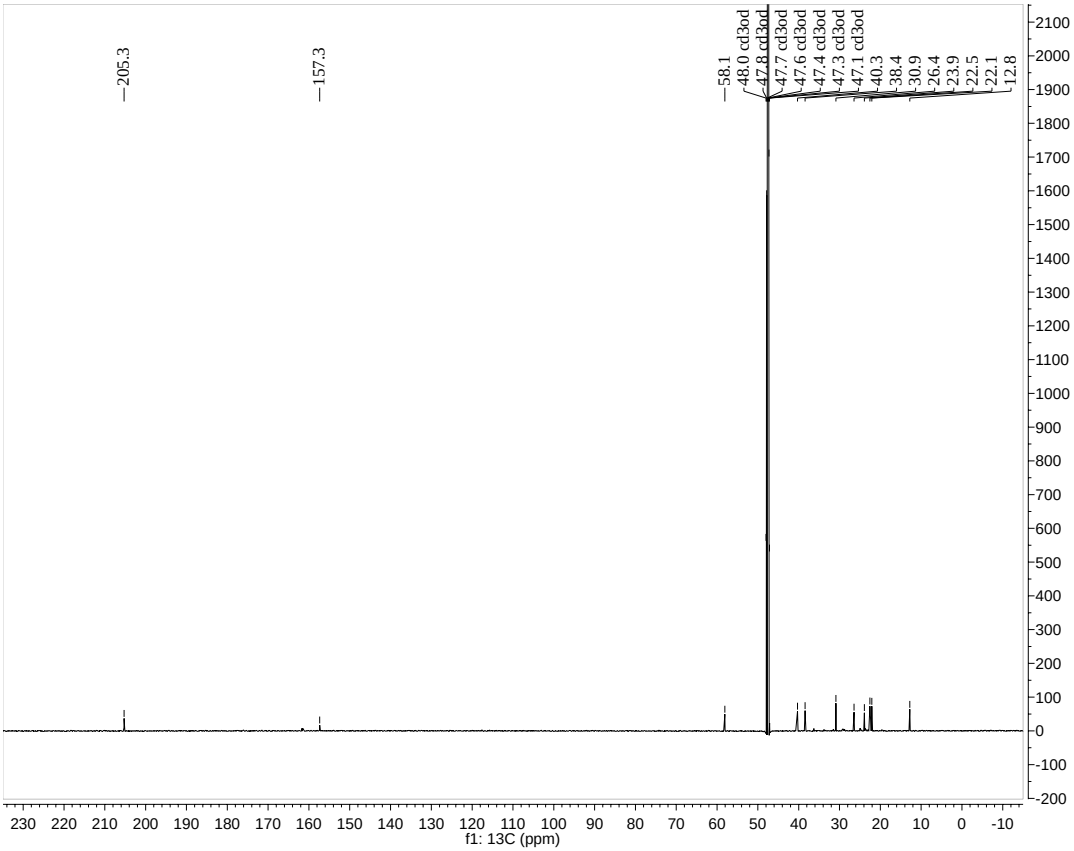
HSQC, highlighted peaks coalesced in ^{13}CNMR

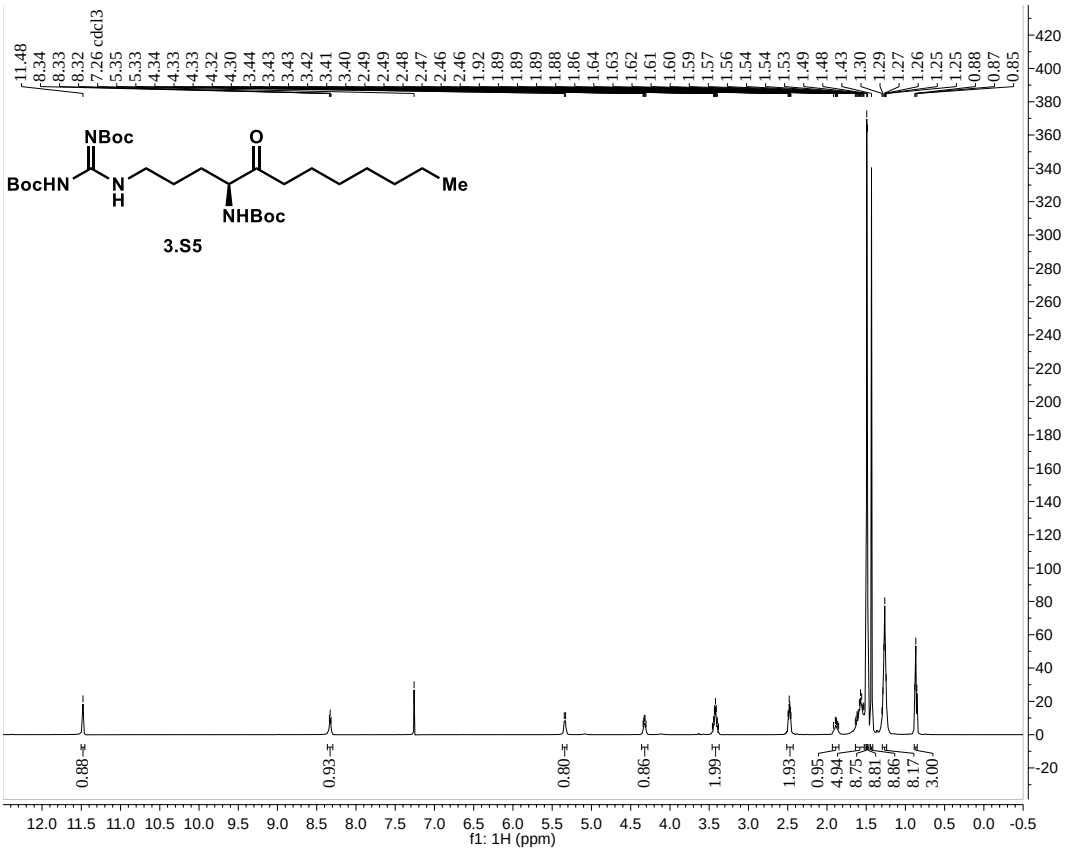


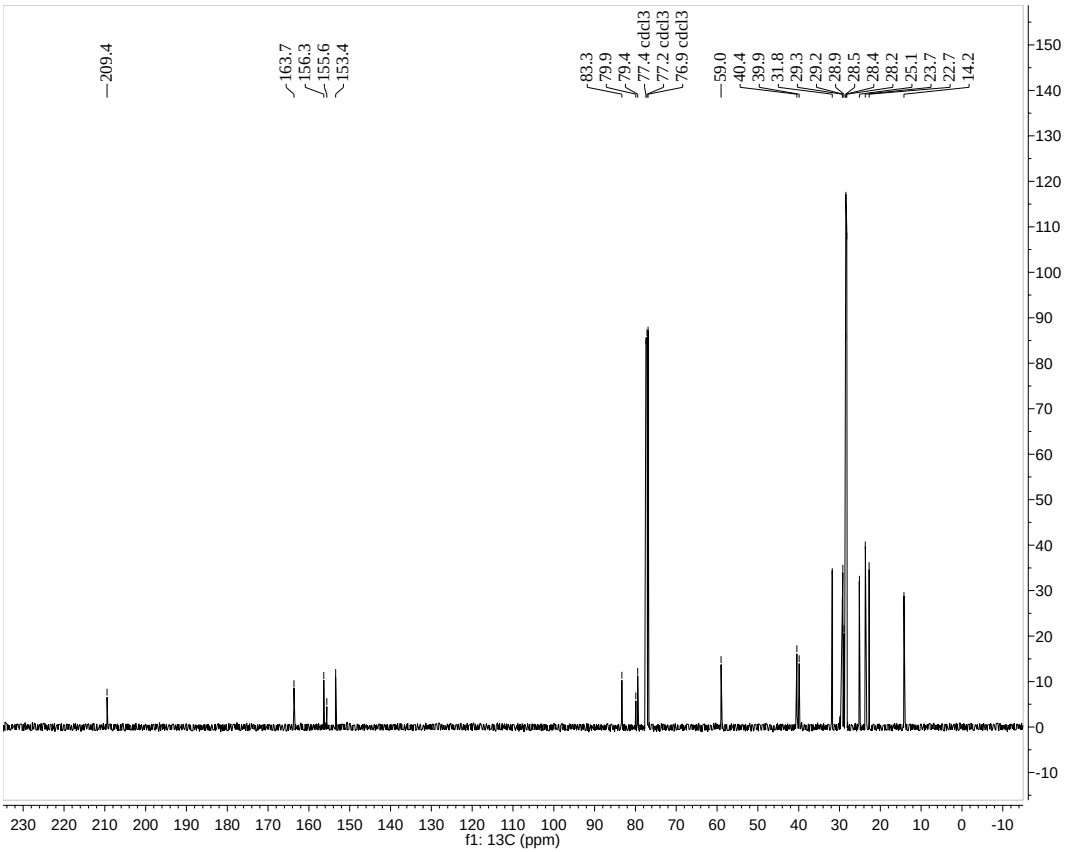

11.48
8.34
8.33
8.32
7.26 cdcl3
5.35
5.33
4.32
3.45
3.45
3.43
3.43
3.41
3.40
3.40
2.50
2.48
2.48
2.46
2.46
1.89
1.61
1.60
1.59
1.57
1.50
1.49
1.44
1.35
1.34
1.34
1.32
1.30
1.29
1.28
1.27
1.25
1.24
1.24
0.90
0.88
0.87
NBoc
O
BocHN
N
H
Me
NHBoc
3.S4
0.77
0.78
0.69
0.77
1.91
1.78
0.84
3.43
8.56
8.66
8.47
5.93
3.00
f1: 1H (ppm)

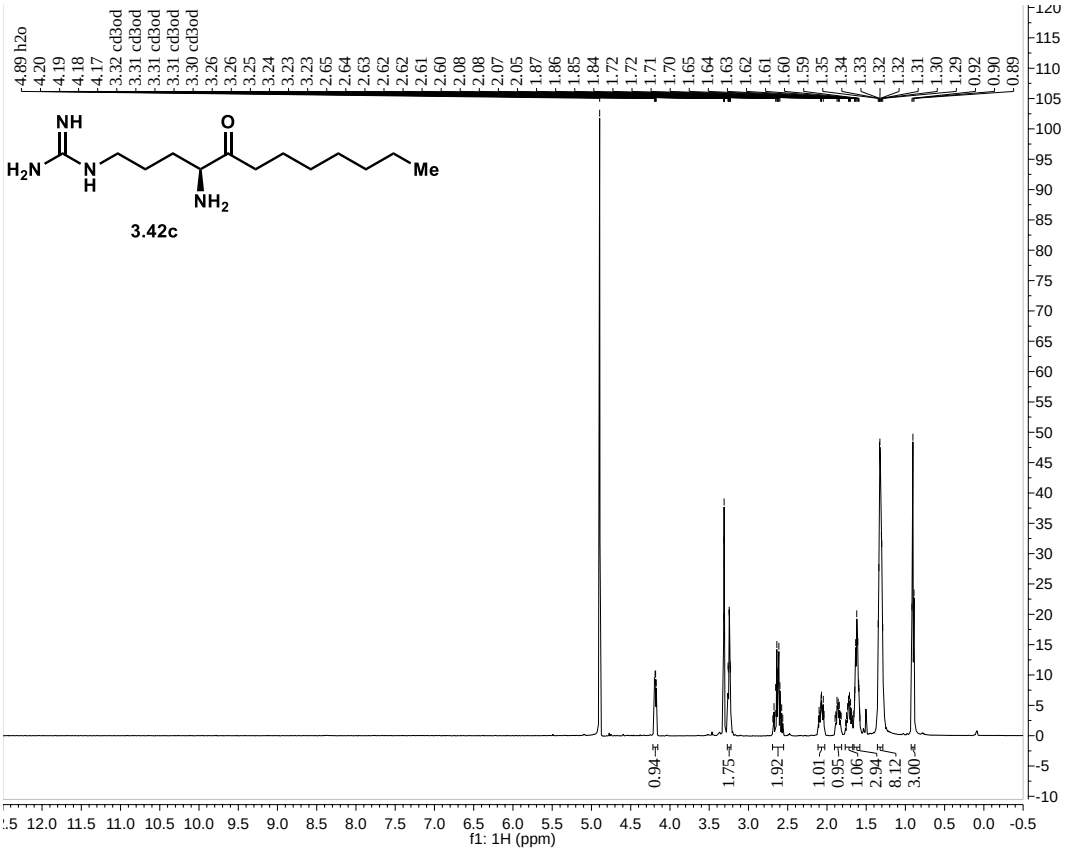

209.4
163.7
156.4
155.6
153.5
83.3
79.9
79.4
77.5 cdcl3
77.2 cdcl3
76.8 cdcl3
59.0
40.4
39.8
31.5
28.9
28.5
28.4
28.2
25.1
23.4
22.6
14.0
f1: 13C (ppm)

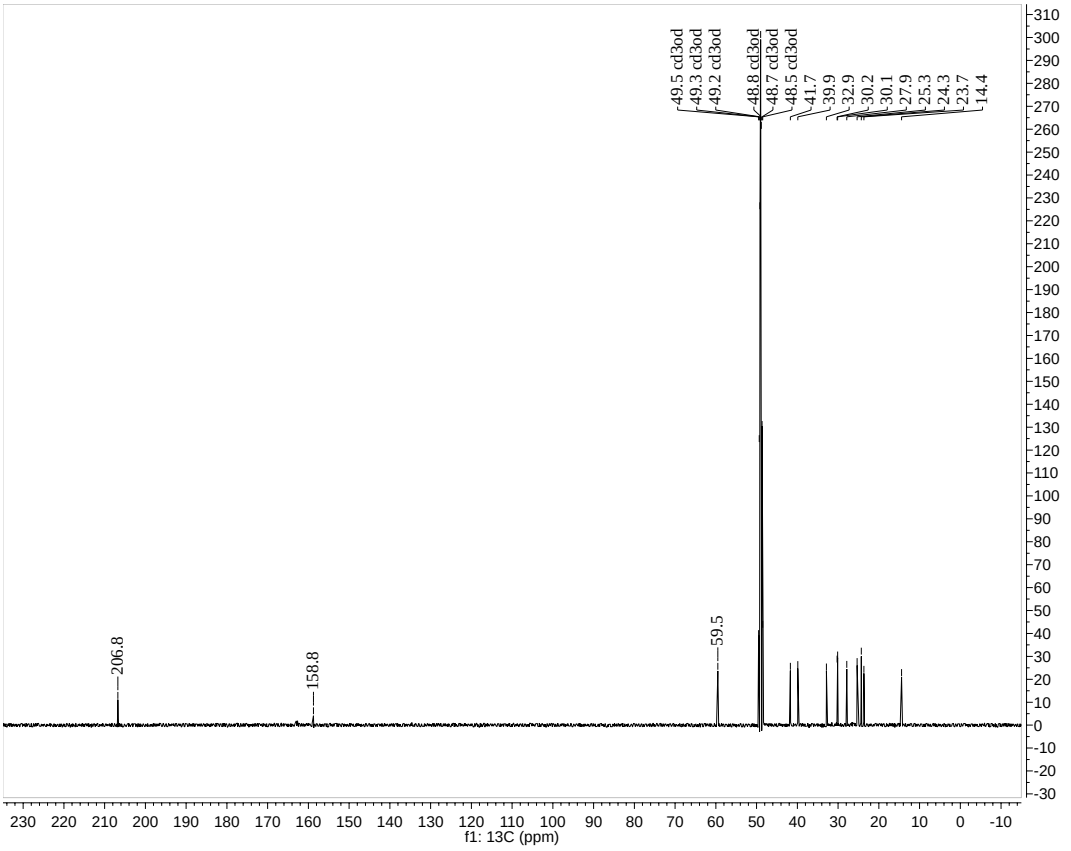

4.84 d2o
4.18
4.18
4.17
4.16
3.31 cd3od
3.30 cd3od
3.30 cd3od
3.30 cd3od
3.29 cd3od
3.25
3.24
3.23
3.23
3.22
2.64
2.63
2.62
2.61
2.60
2.59
2.04
2.03
1.70
1.63
1.63
1.62
1.61
1.61
1.60
1.60
1.60
1.35
1.35
1.35
1.34
1.33
1.33
1.33
1.32
1.32
1.31
1.31
1.31
1.28
0.92
0.92
0.91
0.90
0.90
NH
O
H2N
N
H
Me
NH2
3.42b
0.86
1.60
1.75
1.10
0.88
1.00
2.84
4.05
3.00
f1: 1H (ppm)

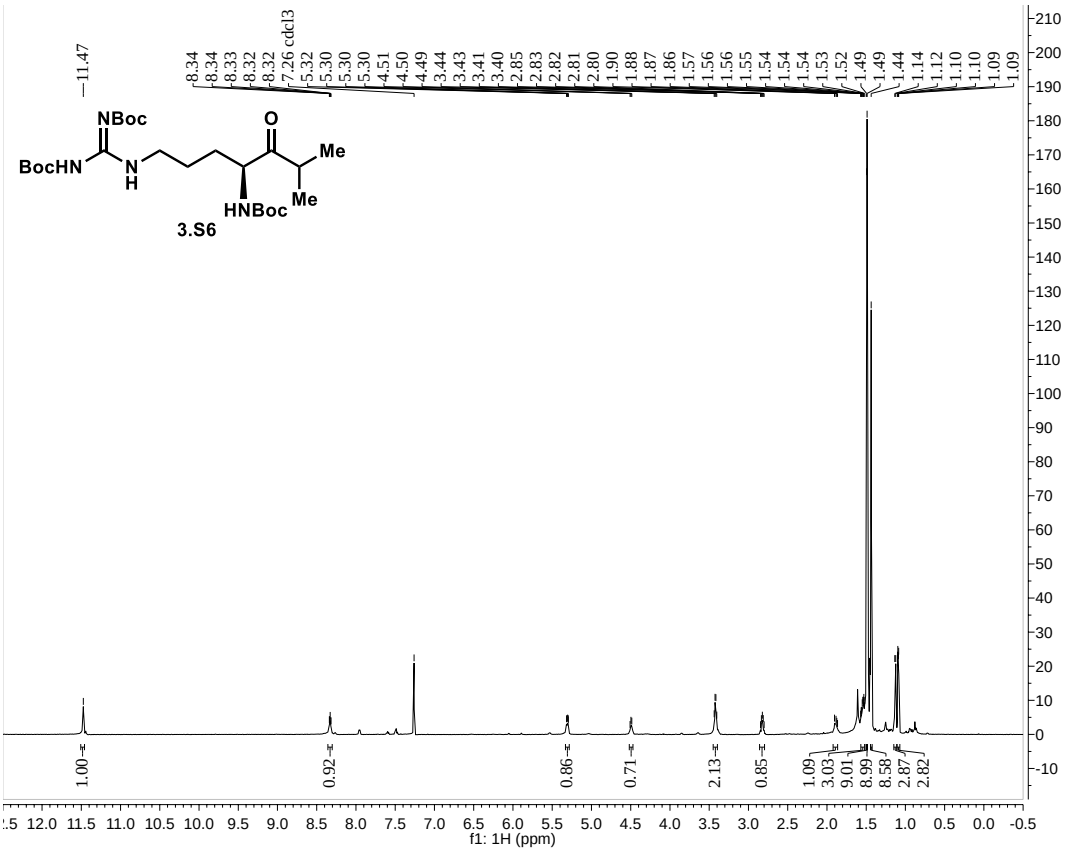

205.3
157.3
58.1
48.0 cd3od
47.8 cd3od
47.7 cd3od
47.6 cd3od
47.4 cd3od
47.3 cd3od
47.1 cd3od
40.3
38.4
30.9
26.4
23.9
22.5
22.1
12.8
f1: 13C (ppm)

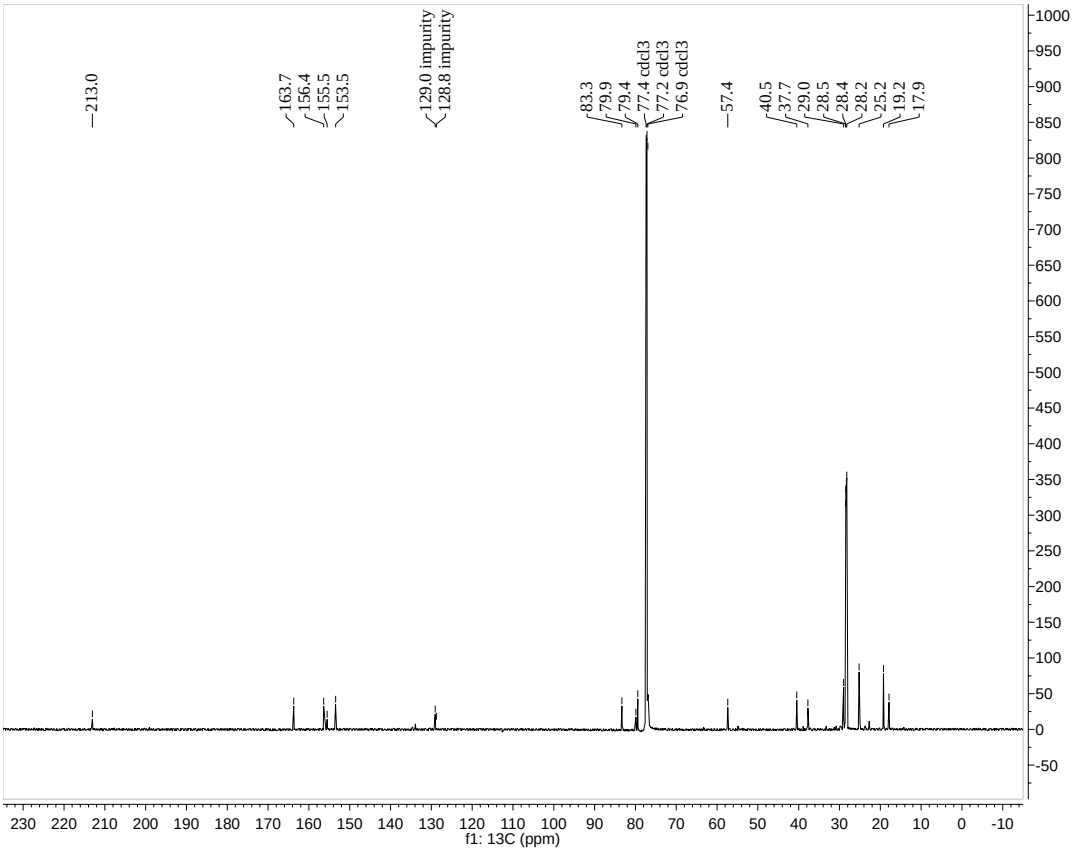

11.48
8.34
8.33
8.32
7.26 cdcl3
5.35
5.33
4.34
4.33
4.33
4.32
4.30
3.44
3.43
3.43
3.42
3.41
3.40
2.49
2.49
2.48
2.47
2.46
2.46
1.92
1.89
1.89
1.89
1.88
1.86
1.64
1.63
1.62
1.61
1.60
1.59
1.57
1.56
1.54
1.54
1.53
1.49
1.48
1.43
1.30
1.29
1.27
1.26
1.25
1.25
0.88
0.87
0.85
NBoc
O
BocHN
N
H
Me
NHBoc
3.S5
0.88
0.93
0.80
0.86
1.99
1.93
0.95
4.94
8.75
8.81
8.86
8.17
3.00
f1: 1H (ppm)

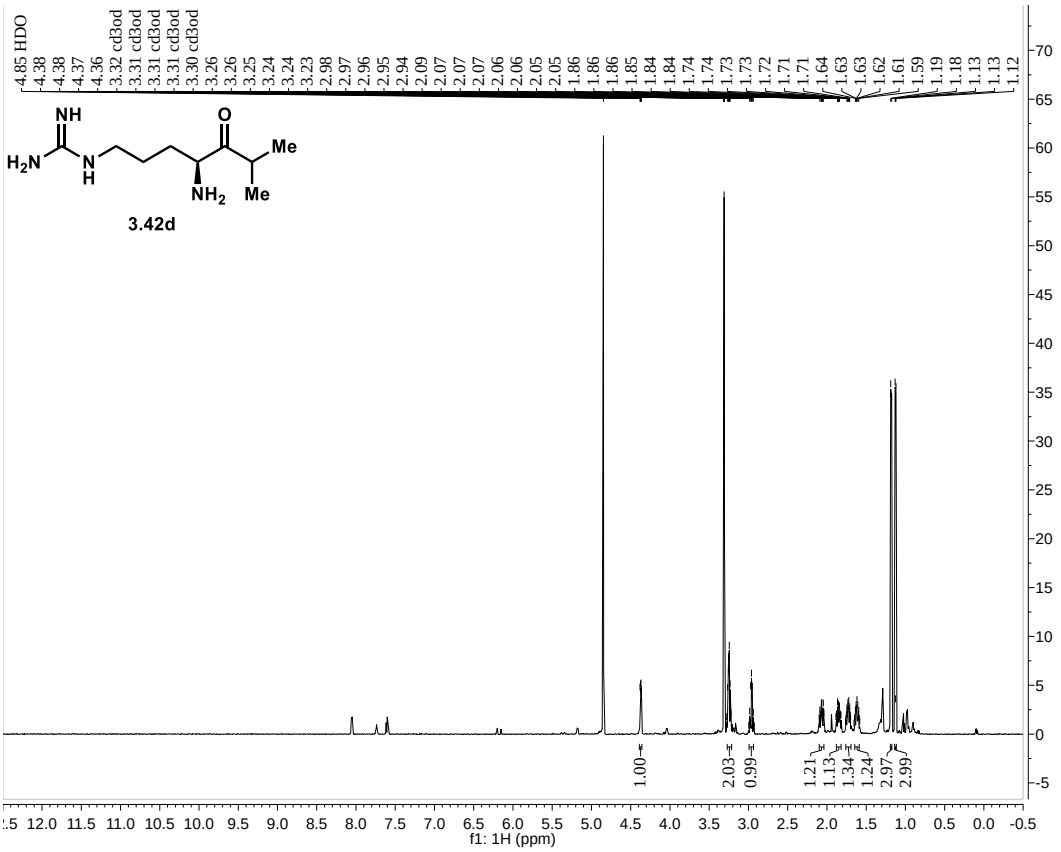

209.4
163.7
156.3
155.6
153.4
83.3
79.9
79.4
77.4 cdcl3
77.2 cdcl3
76.9 cdcl3
59.0
40.4
39.9
31.8
29.3
29.2
28.9
28.5
28.4
28.2
25.1
23.7
22.7
14.2
f1: 13C (ppm)

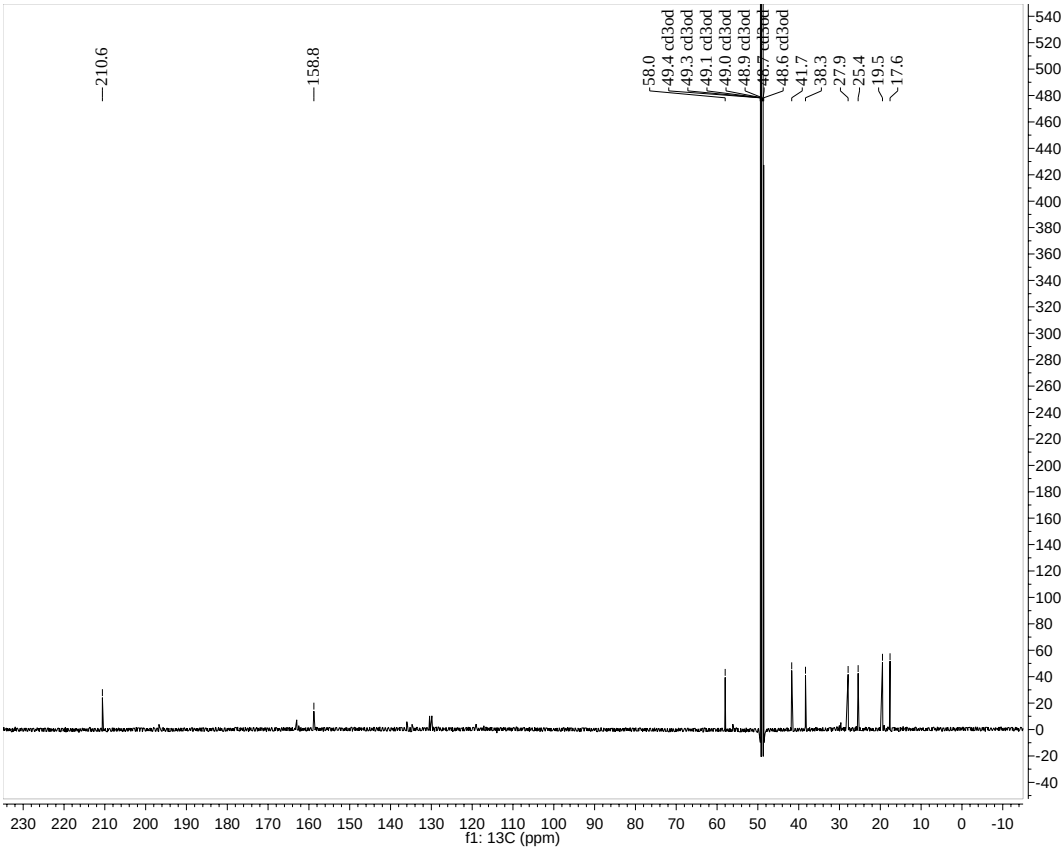

4.89 h2o
4.20
4.19
4.18
4.17
3.32 cd3od
3.31 cd3od
3.31 cd3od
3.31 cd3od
3.30 cd3od
3.26
3.25
3.24
3.23
3.23
2.65
2.64
2.63
2.62
2.62
2.61
2.60
2.08
2.08
2.07
2.05
1.87
1.86
1.85
1.84
1.72
1.72
1.71
1.70
1.65
1.64
1.63
1.62
1.61
1.60
1.59
1.35
1.34
1.33
1.32
1.32
1.31
1.30
1.29
0.92
0.90
0.89
NH
O
H2N
N
H
Me
NH2
3.42c
0.94
1.75
1.92
1.01
0.95
1.06
2.94
8.12
3.00
f1: 1H (ppm)

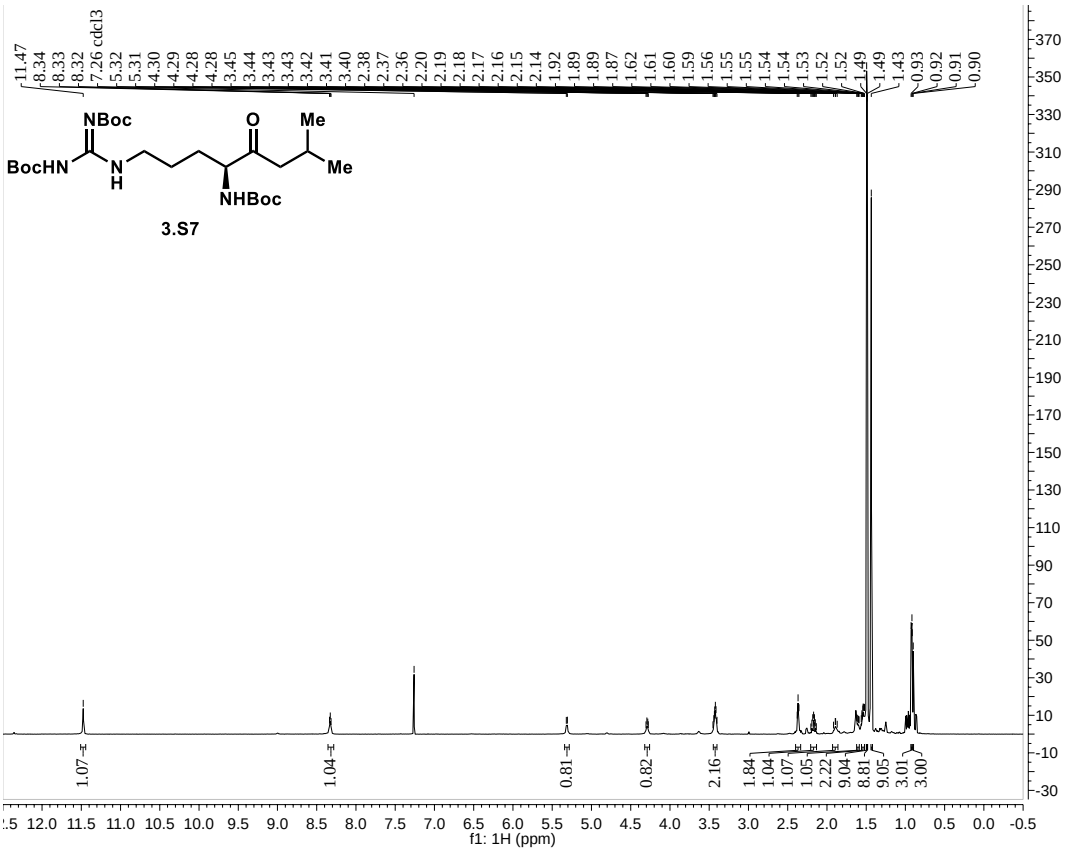

49.5 cd3od
49.3 cd3od
49.2 cd3od
48.8 cd3od
48.7 cd3od
48.5 cd3od
41.7
39.9
32.9
30.2
30.1
27.9
25.3
24.3
23.7
14.4
206.8
158.8
59.5
f1: 13C (ppm)

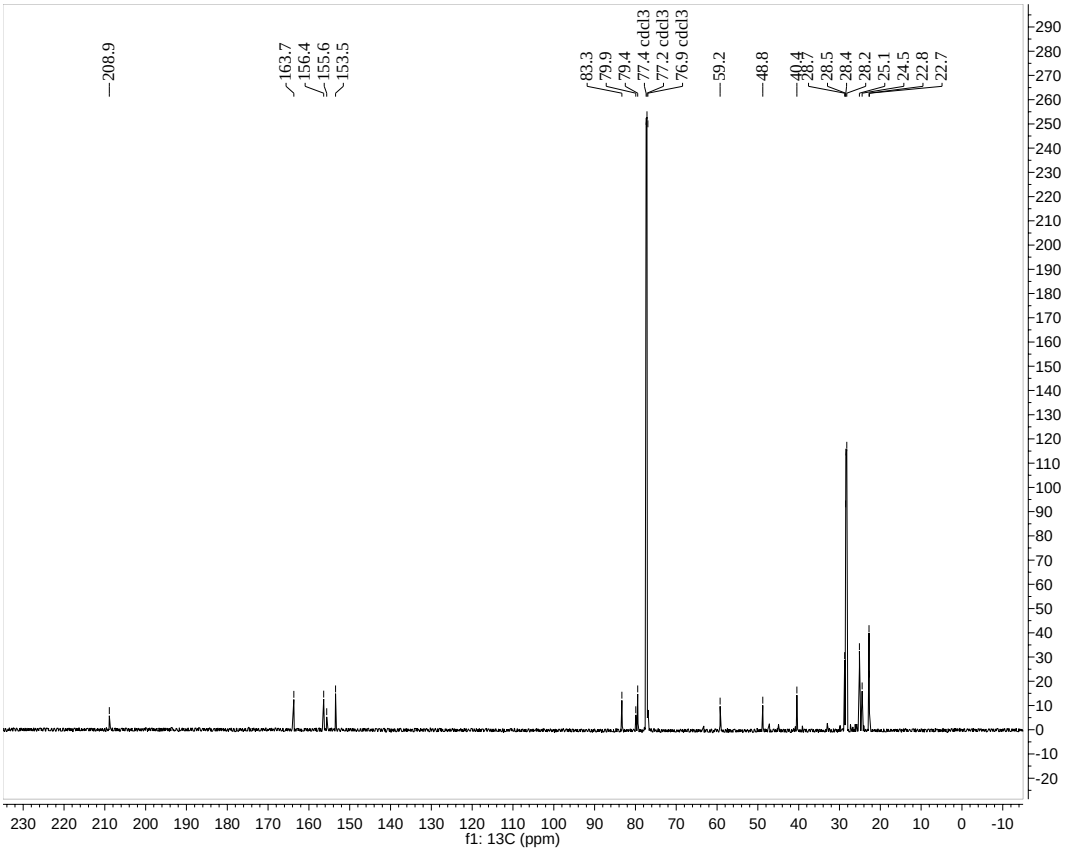

11.47
8.34
8.34
8.33
8.32
8.32
8.32
7.26 cdcl3
5.32
5.30
5.30
5.30
4.51
4.50
4.49
3.44
3.43
3.41
3.40
2.85
2.83
2.82
2.81
2.80
1.90
1.88
1.87
1.86
1.57
1.56
1.56
1.55
1.54
1.54
1.54
1.53
1.52
1.49
1.49
1.44
1.14
1.12
1.10
1.10
1.09
1.09
NBoc
O
BocHN
N
H
Me
Me
HNBoc
3.S6
1.00
0.92
0.86
0.71
2.13
0.85
1.09
3.03
9.01
8.99
8.58
2.87
2.82
f1: 1H (ppm)

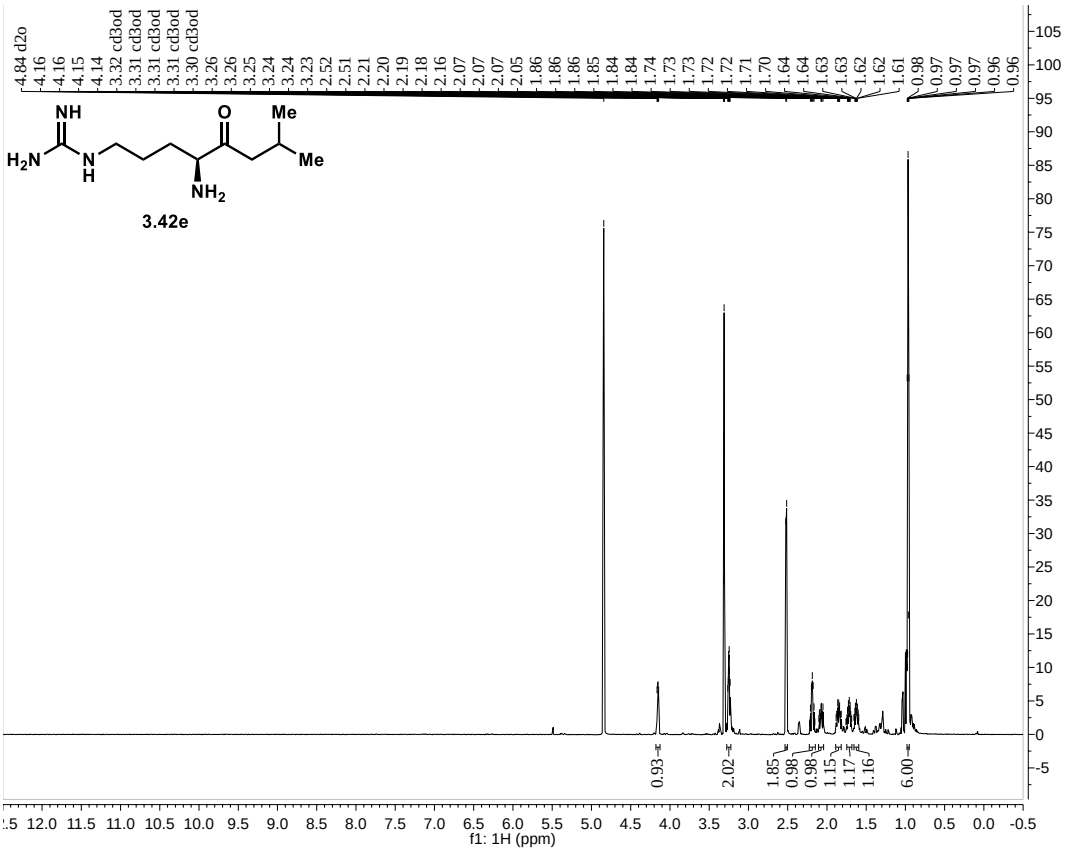

213.0
163.7
156.4
155.5
153.5
129.0 impurity
128.8 impurity
83.3
79.9
79.4
77.4 cdcl3
77.2 cdcl3
76.9 cdcl3
57.4
40.5
37.7
29.0
28.5
28.4
28.2
25.2
19.2
17.9
f1: 13C (ppm)

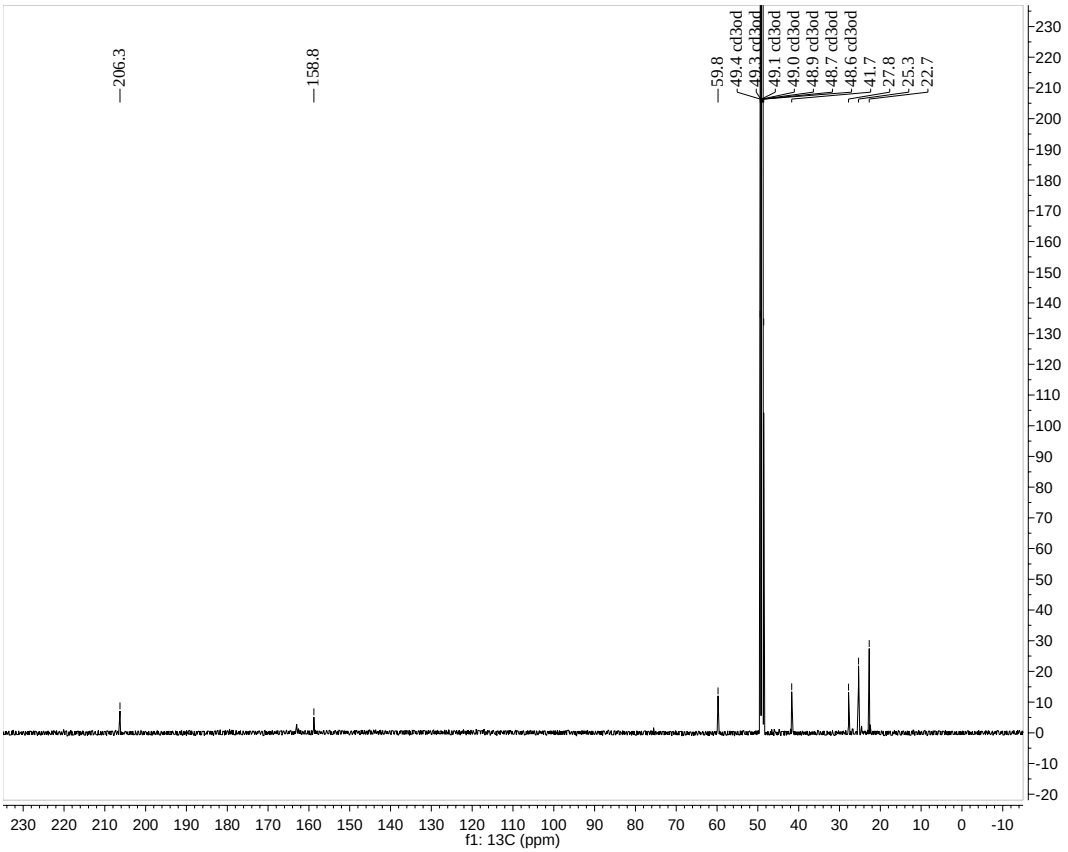

4.85 HDO
4.38
4.38
4.37
4.36
3.32 cd3od
3.31 cd3od
3.31 cd3od
3.31 cd3od
3.30 cd3od
3.26
3.25
3.24
3.24
3.23
2.98
2.97
2.96
2.95
2.94
2.09
2.07
2.07
2.07
2.06
2.06
2.05
2.05
1.86
1.86
1.86
1.85
1.84
1.84
1.74
1.74
1.73
1.73
1.72
1.71
1.71
1.64
1.63
1.63
1.62
1.61
1.59
1.19
1.18
1.13
1.12
NH
O
H2N
N
H
Me
NH2
Me
3.42d
1.00
2.03
0.99
1.21
1.13
1.34
1.24
2.97
2.99
f1: 1H (ppm)

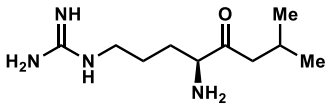

210.6
158.8
58.0
49.4 cd3od
49.3 cd3od
49.1 cd3od
49.0 cd3od
48.9 cd3od
48.7 cd3od
48.6 cd3od
41.7
38.3
27.9
25.4
19.5
17.6
f1: 13C (ppm)


NBoc
O
Me
BocHN
N
H
Me
NHBoc
3.S7
f1: 1H (ppm)


f1: 13C (ppm)

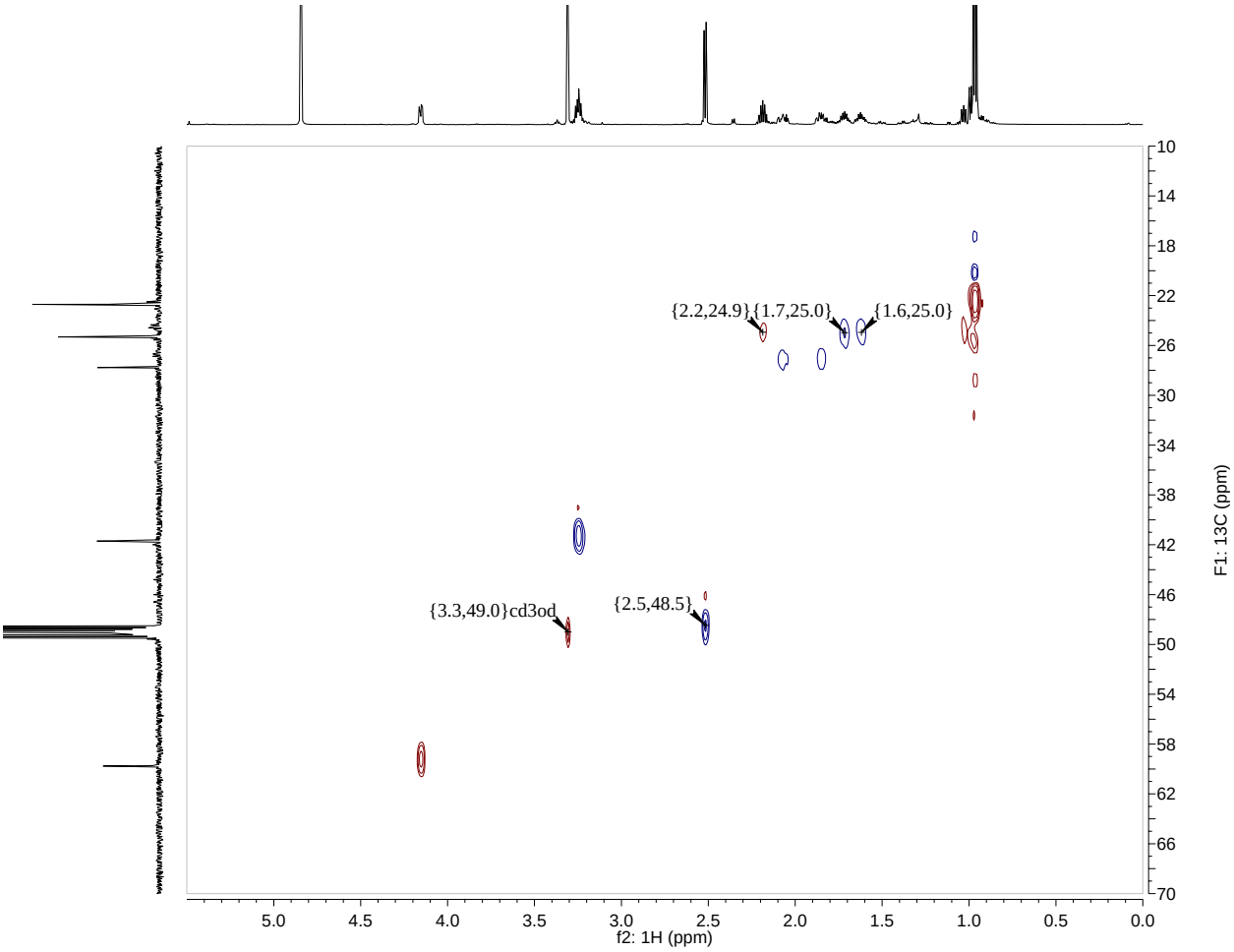

4.84 d2o
4.16
4.16
4.15
4.14
3.32 cd3od
3.31 cd3od
3.31 cd3od
3.31 cd3od
3.30 cd3od
3.26
3.25
3.24
3.24
3.23
2.52
2.51
2.21
2.20
2.19
2.18
2.16
2.07
2.07
2.07
2.05
1.86
1.86
1.85
1.84
1.84
1.74
1.73
1.73
1.72
1.72
1.71
1.70
1.64
1.64
1.63
1.63
1.62
1.62
1.61
0.98
0.97
0.97
0.97
0.96
0.96
NH
O
Me
H2N
N
H
Me
NH2
3.42e
0.93
2.02
1.85
0.98
0.98
1.15
1.17
1.16
6.00
f1: 1H (ppm)

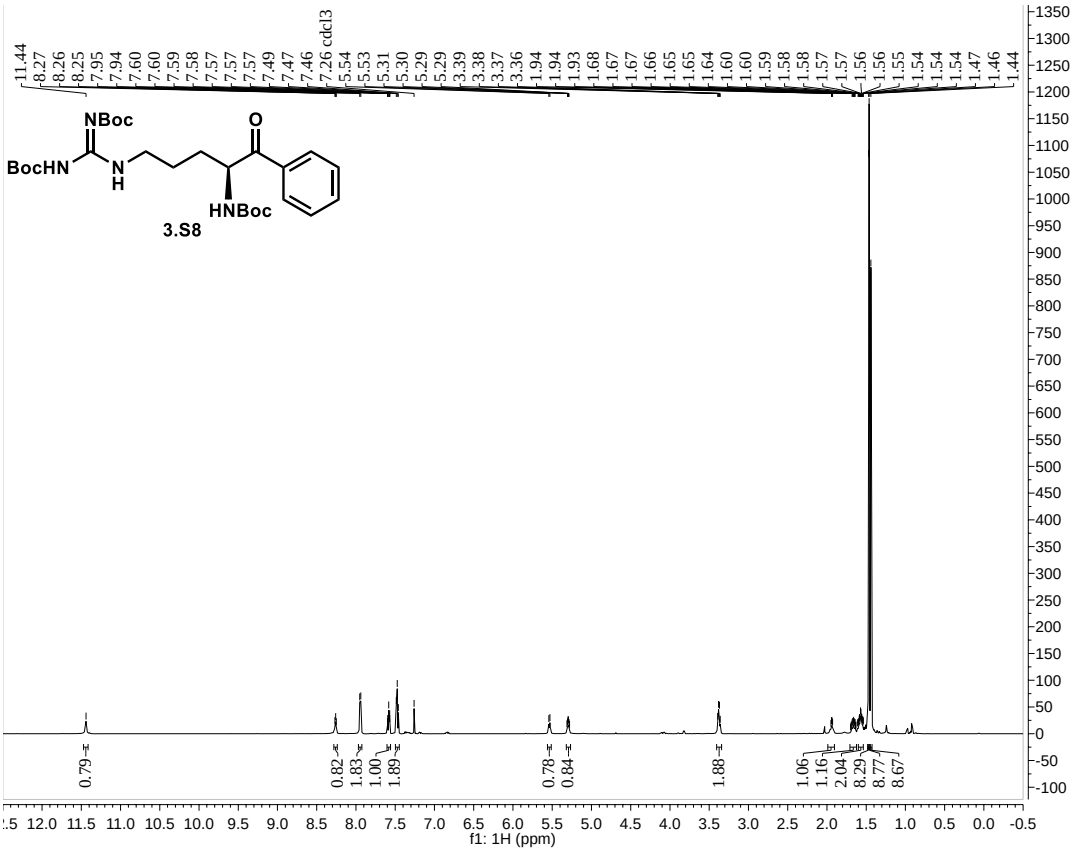

206.3
158.8
59.8
49.4 cd3od
49.3 cd3od
49.1 cd3od
49.0 cd3od
48.9 cd3od
48.7 cd3od
48.6 cd3od
41.7
27.8
25.3
22.7
f1: 13C (ppm)

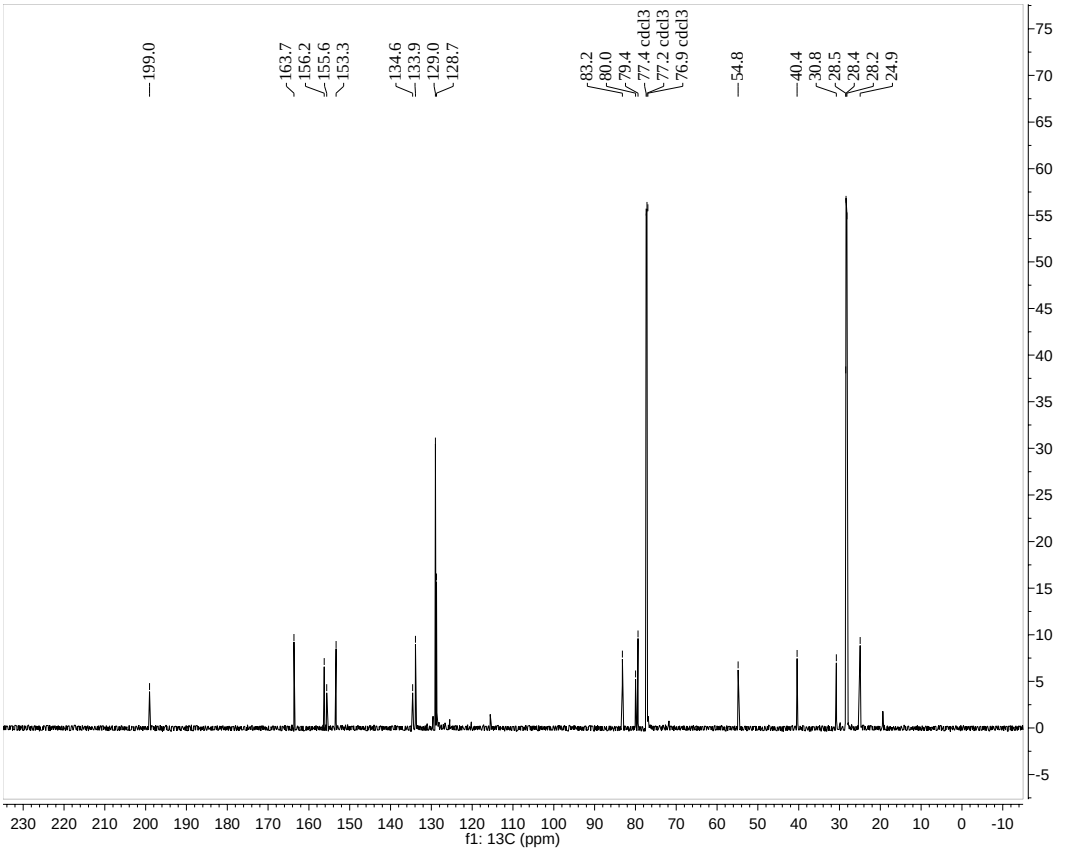


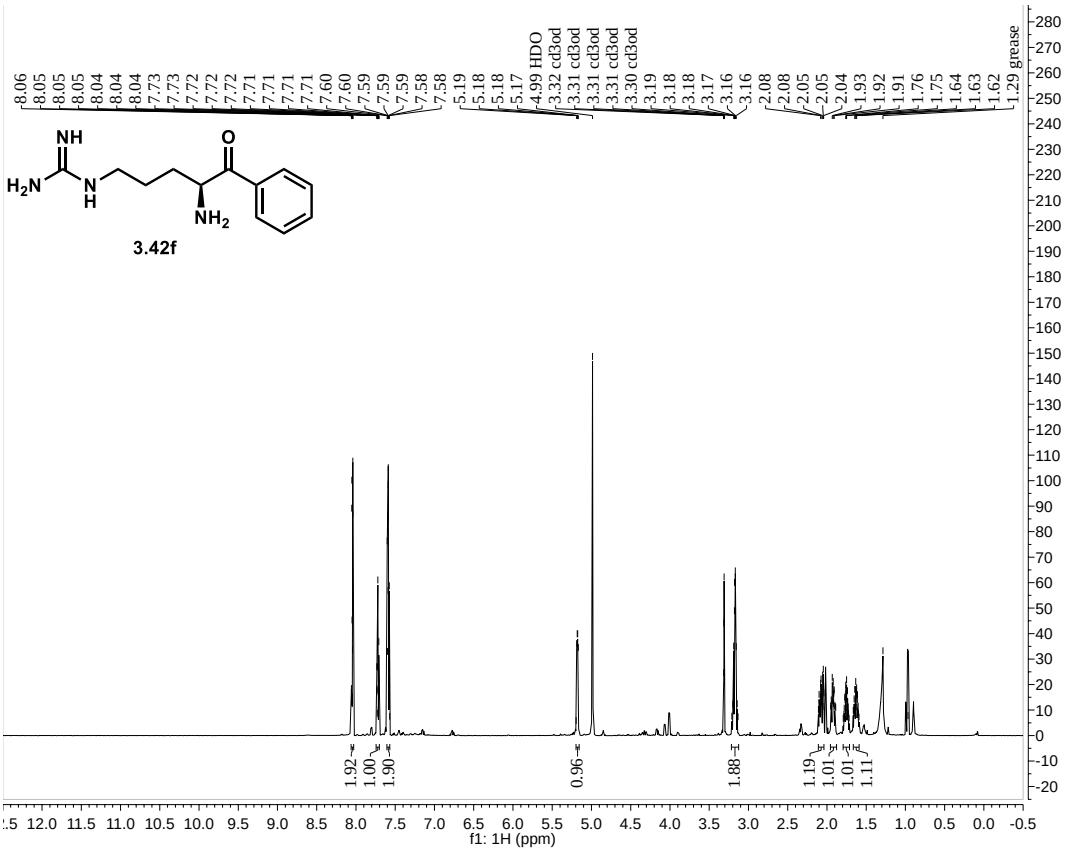
3.42e

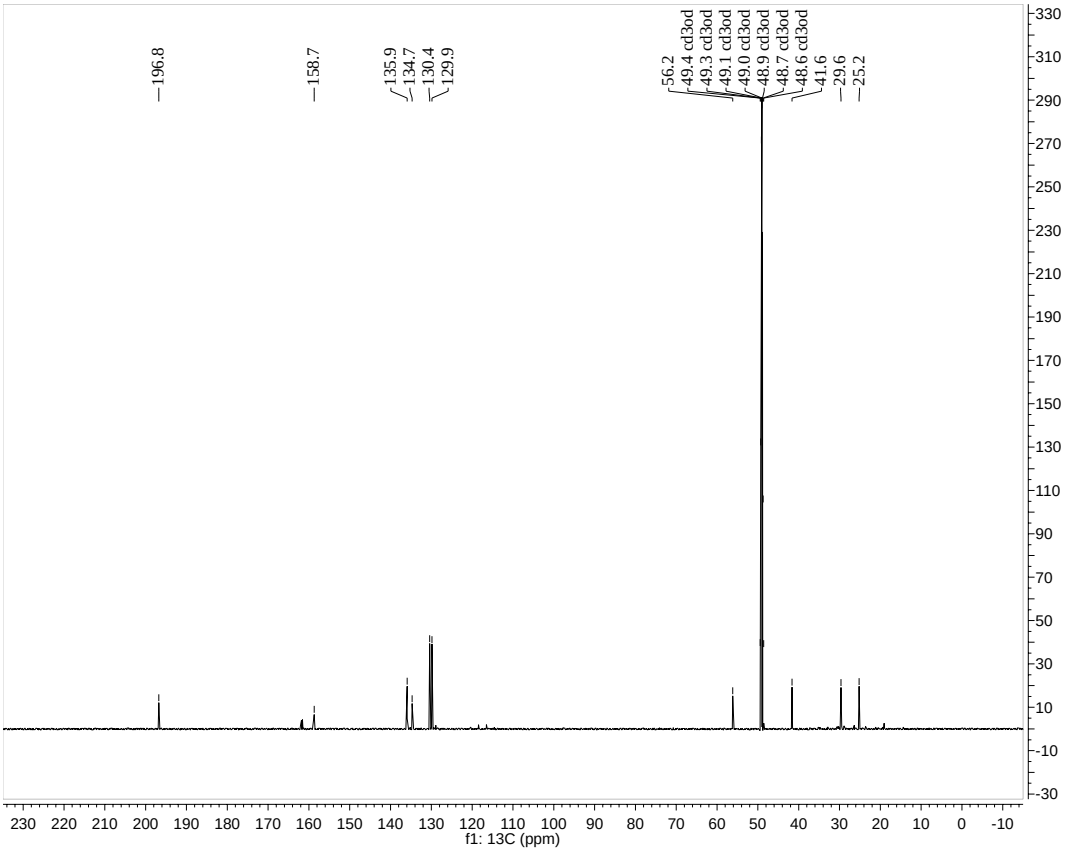
HSQC, highlighted peaks coalesced in ^{13}CNMR

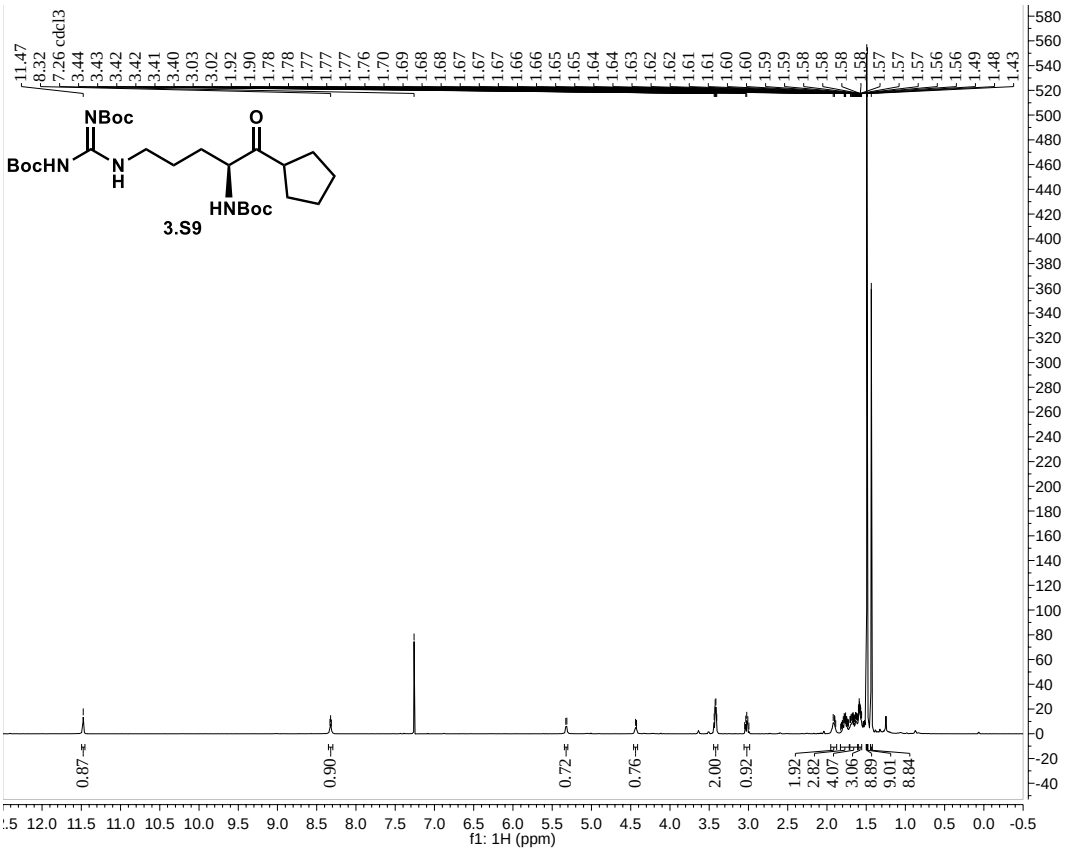


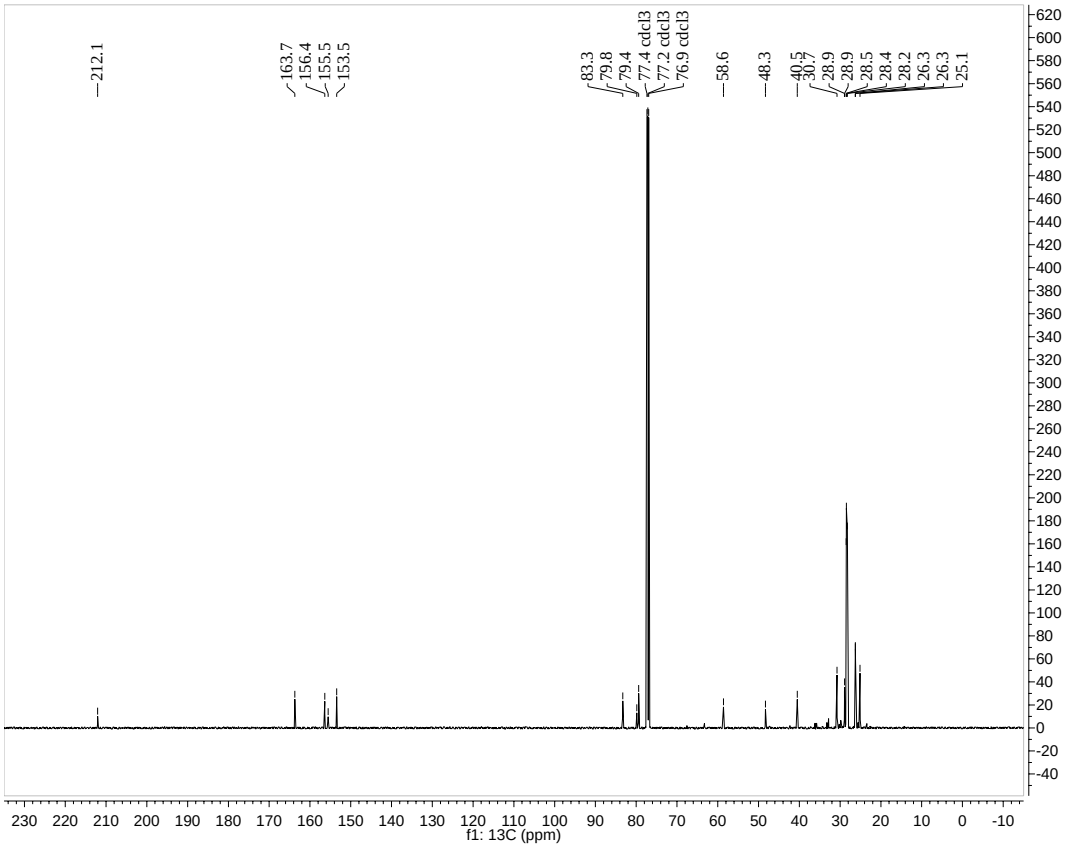

NBoc
BocHN
N
H
O
HNBoc
3.S8
f1: 1H (ppm)

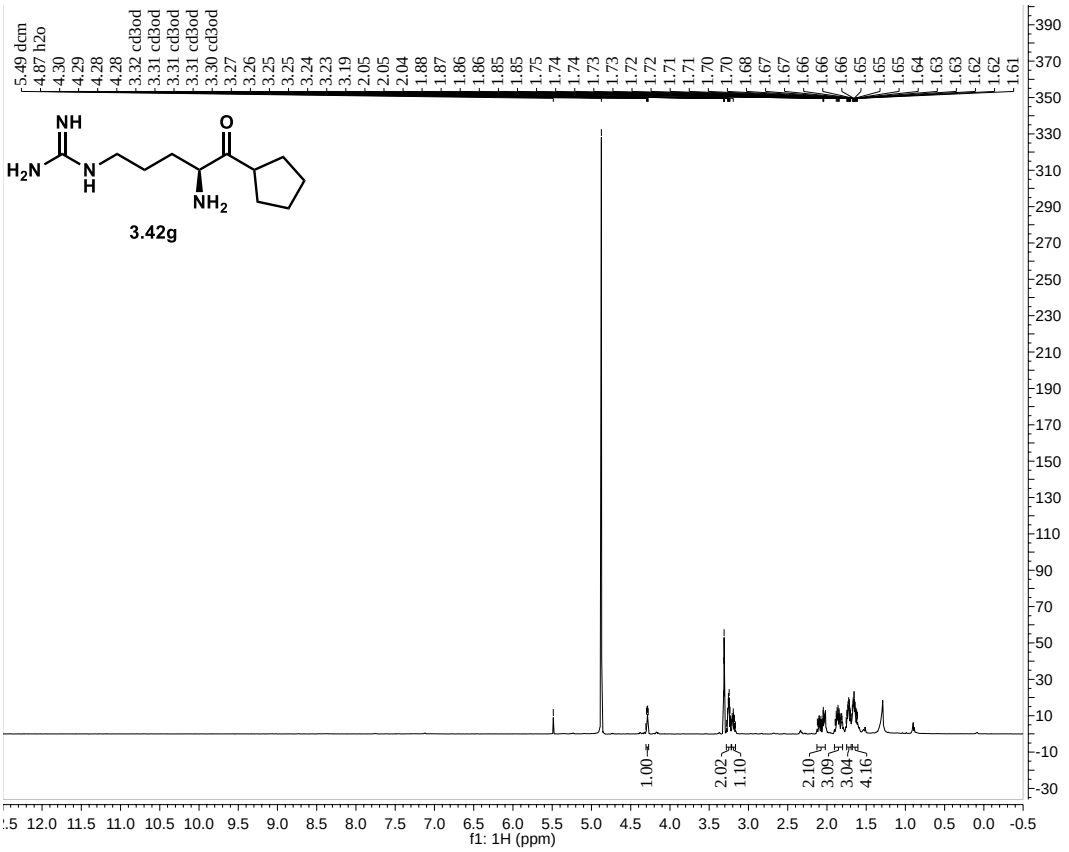

f1: 13C (ppm)

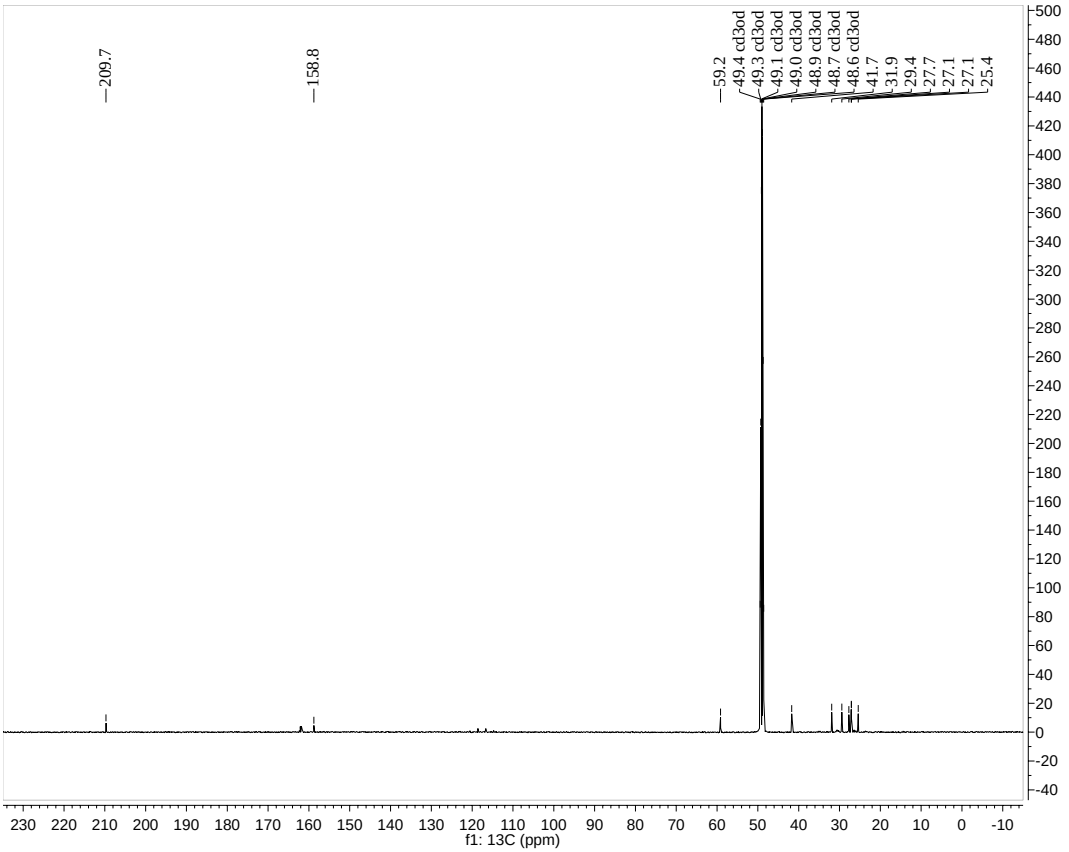

8.06 8.05 8.05 8.05 8.04 8.04 8.04 7.73 7.73 7.72 7.72 7.72 7.71 7.71 7.71 7.71 7.60 7.59 7.59 7.59 7.58 7.58
5.19 5.18 5.18 5.17 4.99 HDO 3.32 cd3od 3.31 cd3od 3.31 cd3od 3.31 cd3od 3.30 cd3od 3.19 3.18 3.18 3.17 3.16 3.16
2.08 2.08 2.05 2.05 2.04 1.93 1.92 1.91 1.76 1.75 1.64 1.63 1.62 1.29 grease
NH
O
H2N
N
H
NH2
3.42f
1.92 1.00 1.90 0.96 1.88 1.19 1.01 1.01 1.11
f1: 1H (ppm)

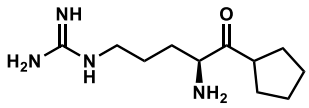

196.8
158.7
135.9
134.7
130.4
129.9
56.2
49.4 cd3od
49.3 cd3od
49.1 cd3od
49.0 cd3od
48.9 cd3od
48.7 cd3od
48.6 cd3od
41.6
29.6
25.2
f1: 13C (ppm)


NBoc
O
BocHN
N
H
HNBoc
3.S9
f1: 1H (ppm)


f1: 13C (ppm)

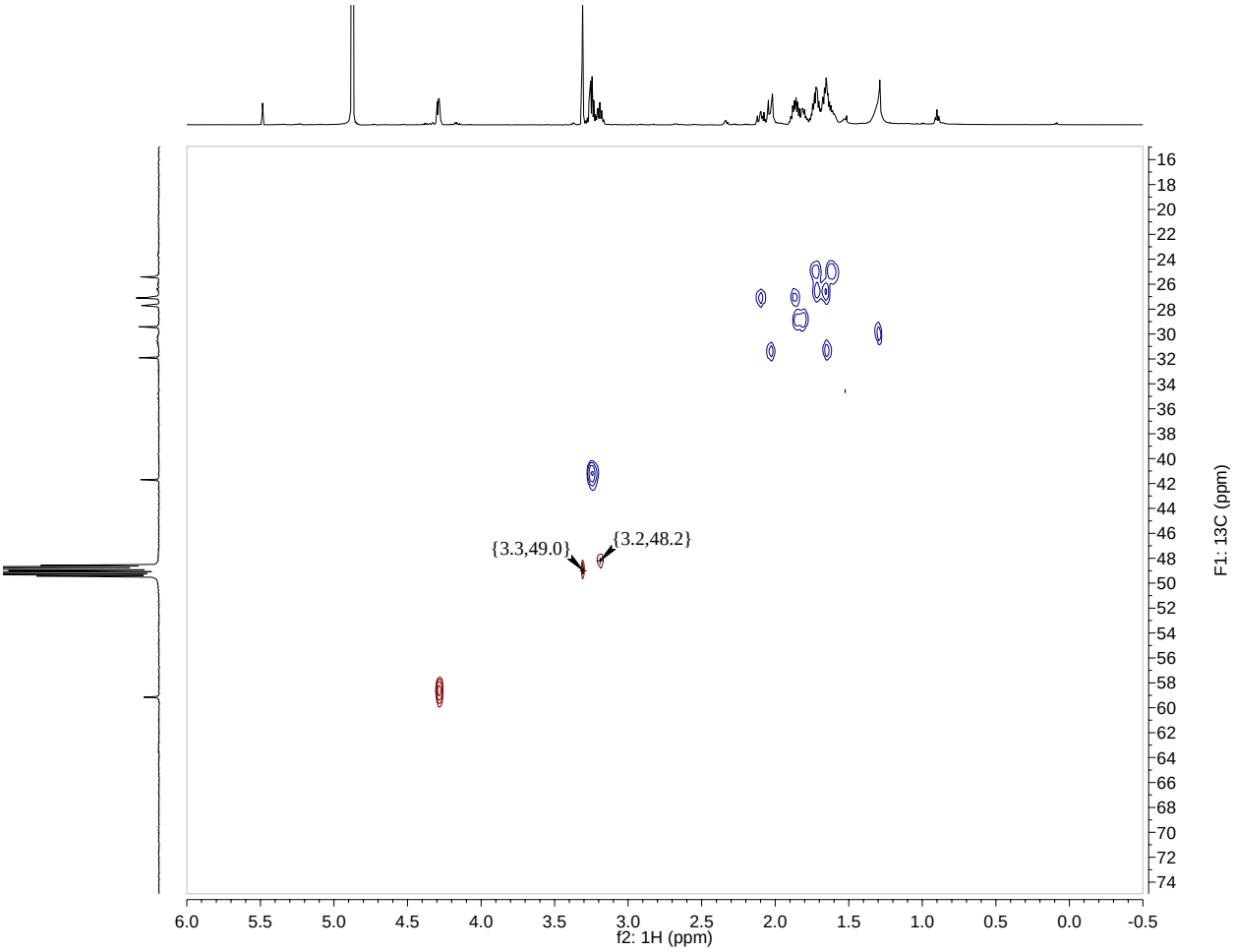

5.49 dcm
4.87 h2o
4.30
4.29
4.28
4.28
3.32 cd3od
3.31 cd3od
3.31 cd3od
3.31 cd3od
3.30 cd3od
3.27
3.26
3.25
3.25
3.24
3.23
3.19
2.05
2.05
2.04
1.88
1.87
1.86
1.86
1.85
1.75
1.74
1.73
1.73
1.72
1.72
1.71
1.71
1.70
1.70
1.68
1.67
1.67
1.66
1.66
1.66
1.65
1.65
1.65
1.64
1.63
1.63
1.62
1.62
1.61
NH
O
H2N
N
H
NH2
3.42g
1.00
2.02
1.10
2.10
3.09
3.04
4.16
f1: 1H (ppm)

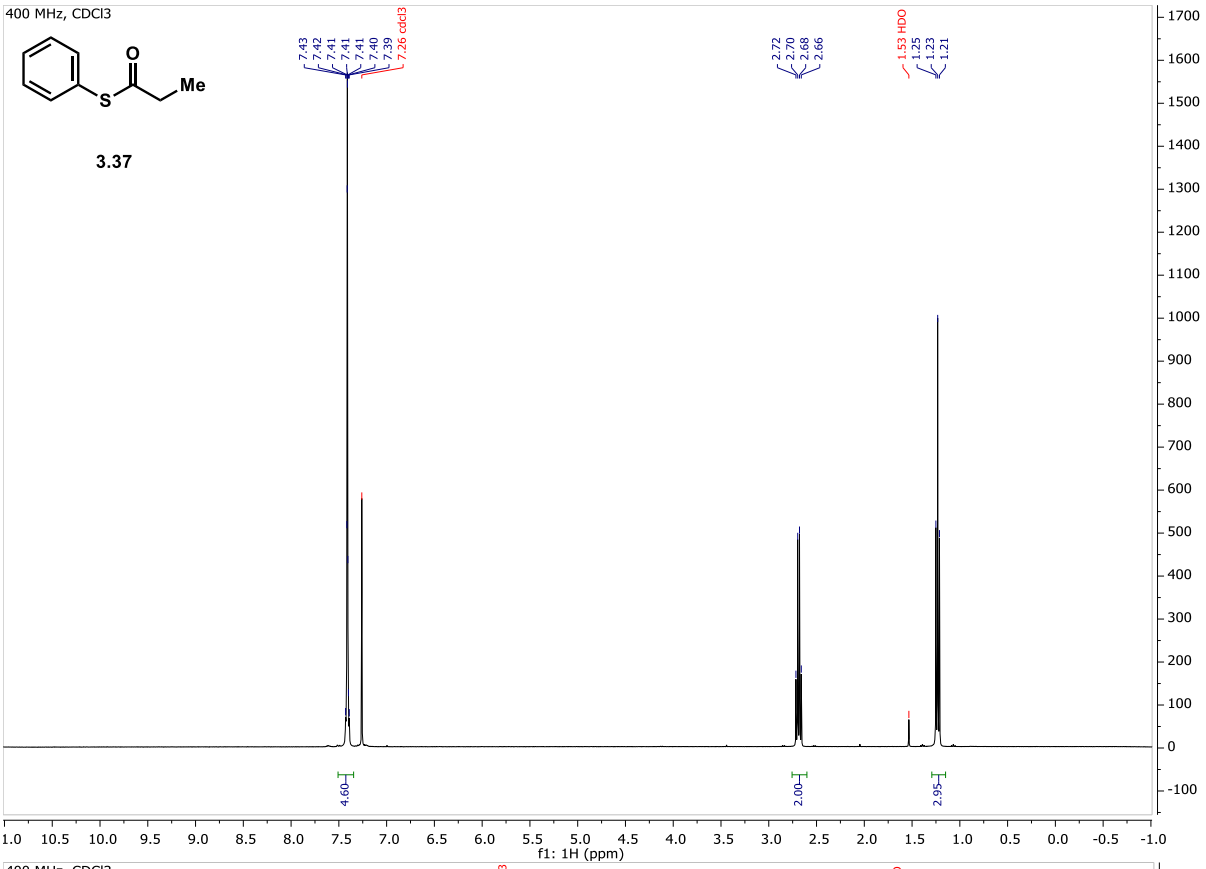

209.7
158.8
59.2
49.4 cd3od
49.3 cd3od
49.1 cd3od
49.0 cd3od
48.9 cd3od
48.7 cd3od
48.6 cd3od
41.7
31.9
29.4
27.7
27.1
27.1
25.4
f1: 13C (ppm)

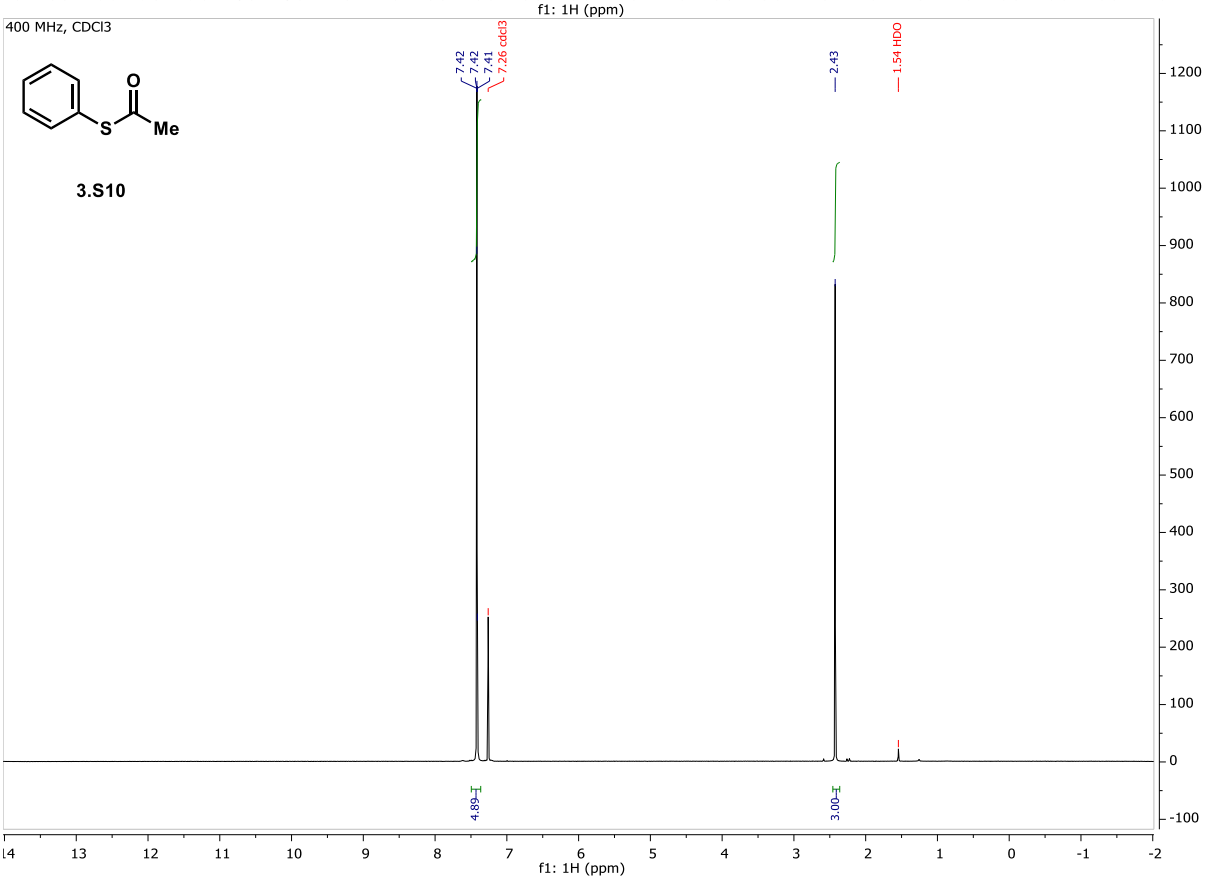


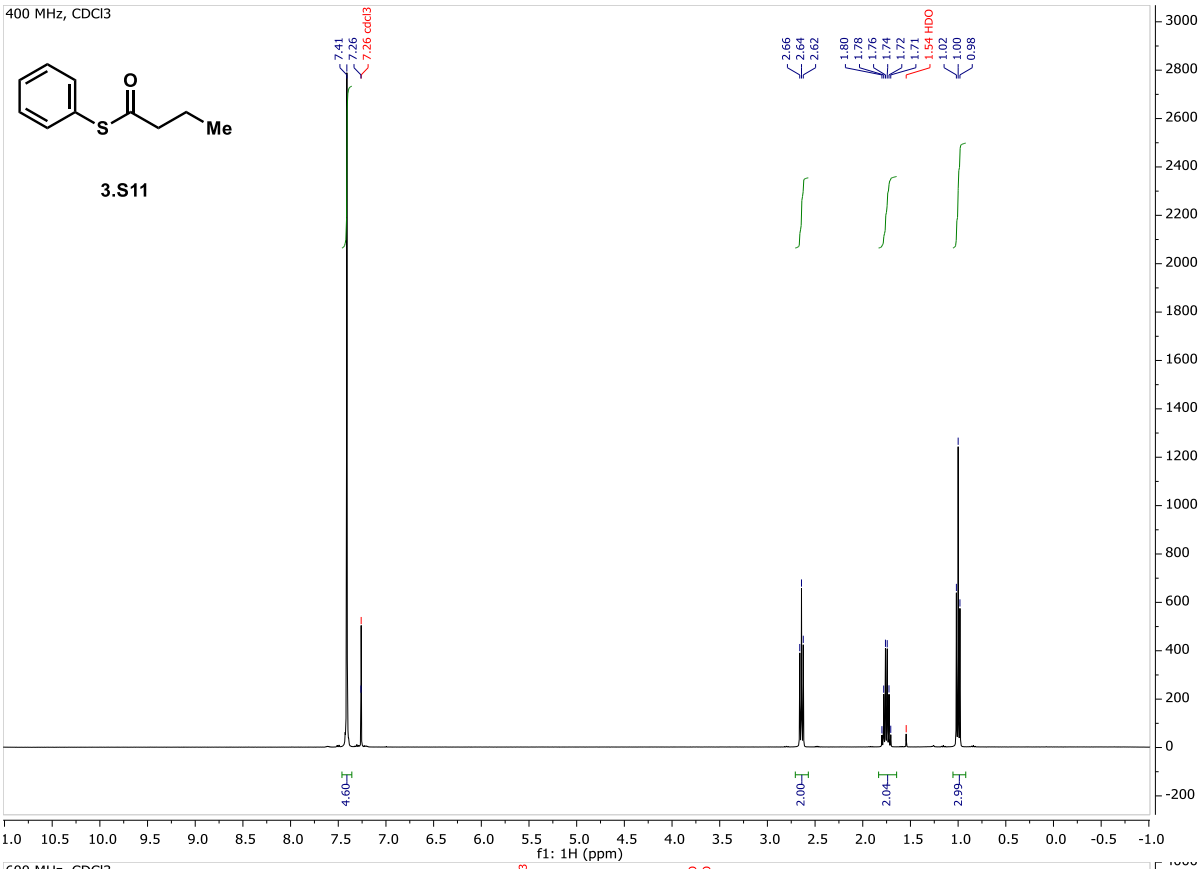
3.42g

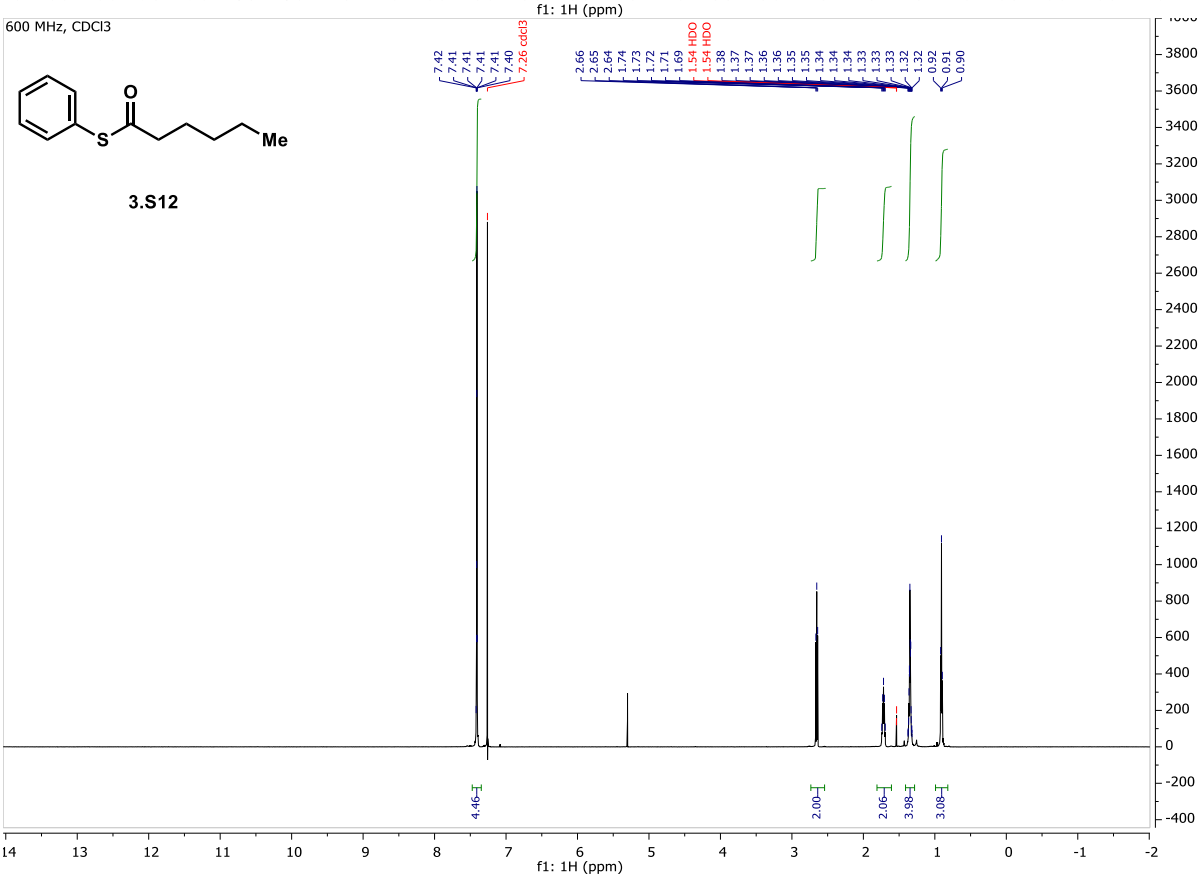
HSQC, highlighted peaks coalesced in ^{13}CNMR

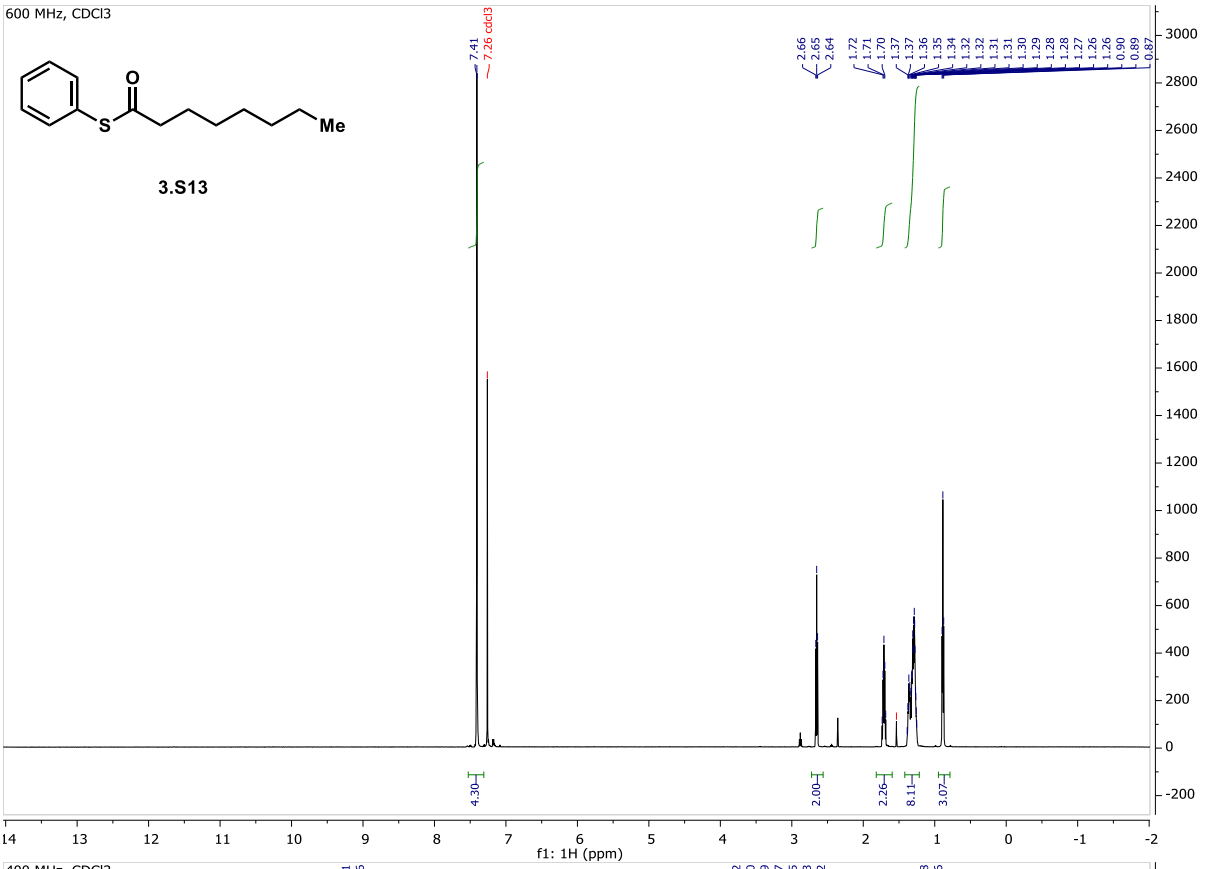


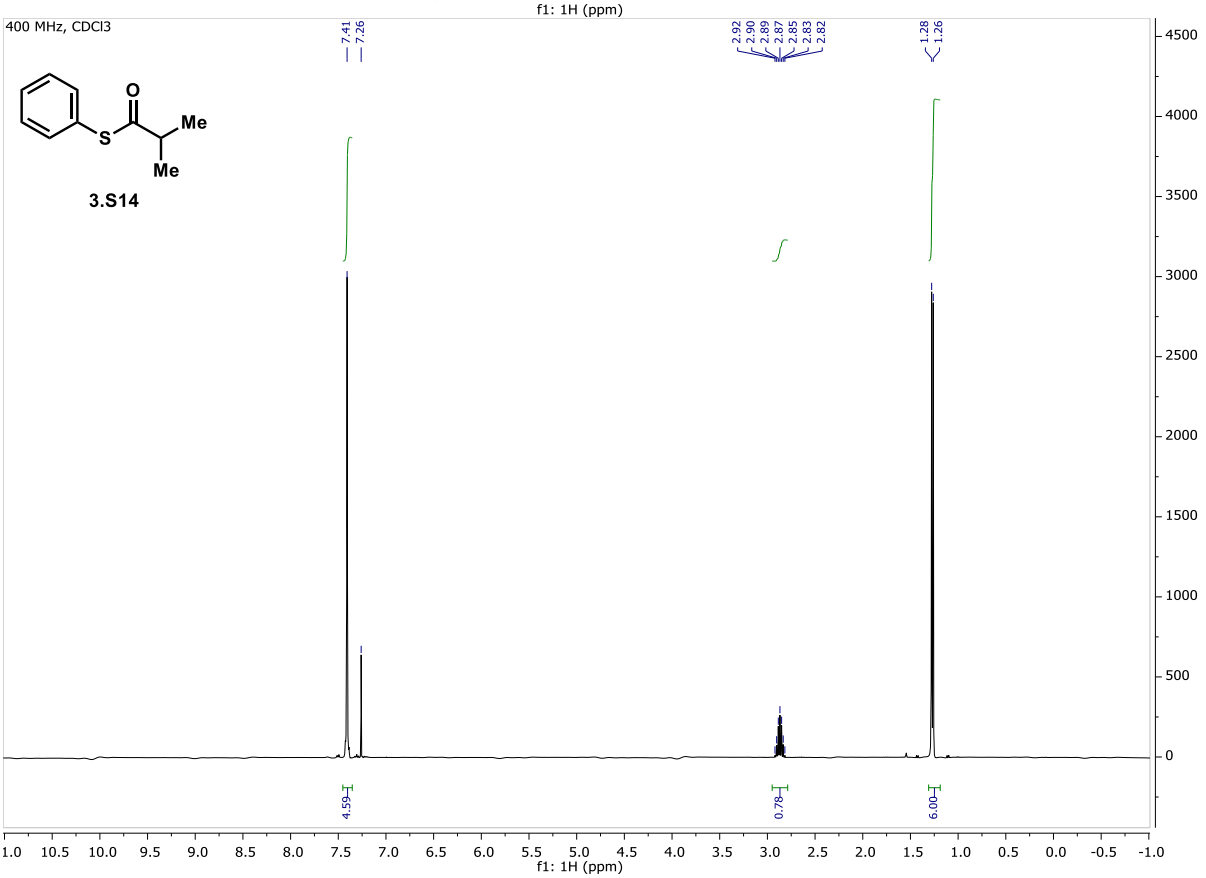

400 MHz, CDCl3
3.37
7.43
7.42
7.41
7.41
7.41
7.40
7.39
7.26 cdcl3
2.72
2.70
2.68
2.66
1.53 HDO
1.25
1.23
1.21
4.60
2.00
2.95
f1: 1H (ppm)

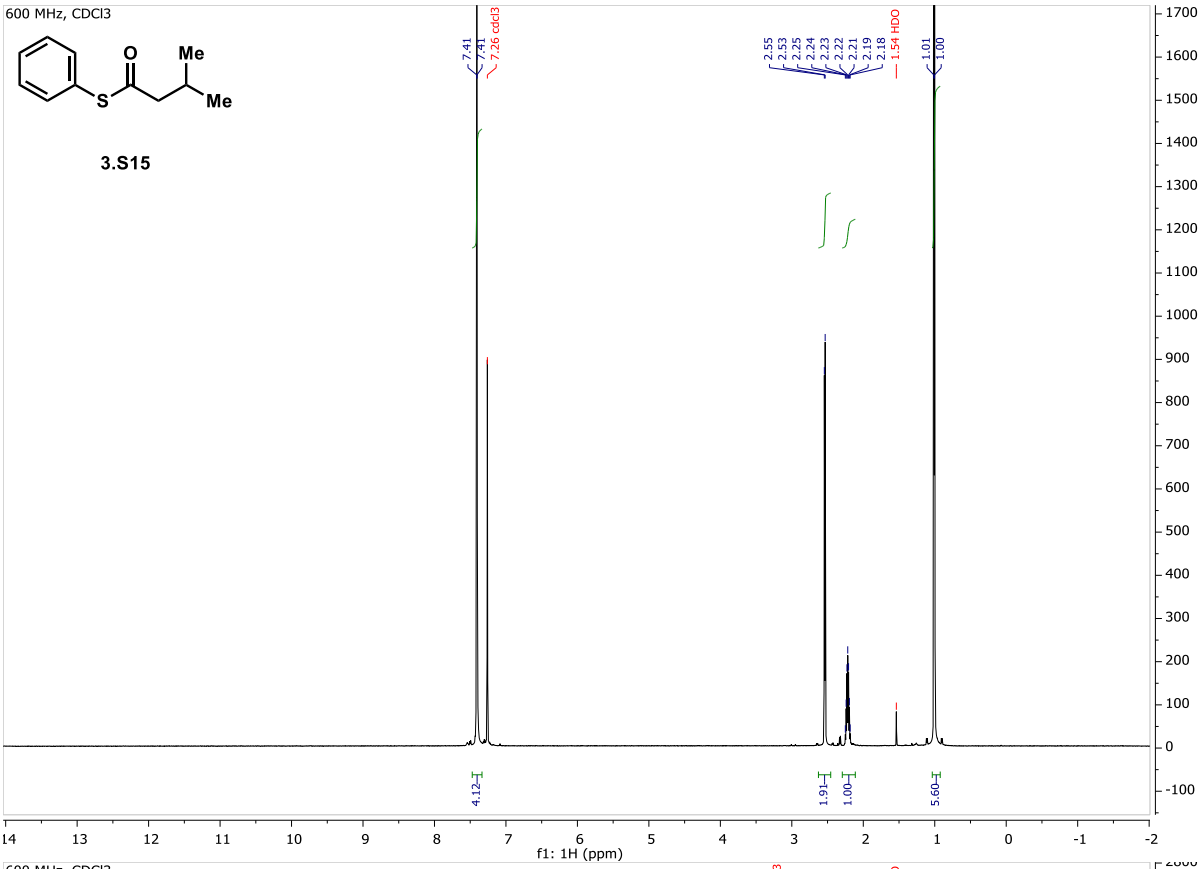

400 MHz, CDCl3
3.S10
7.42
7.42
7.41
7.26 cdcl3
2.43
1.54 HDO
4.89
3.00
f1: 1H (ppm)

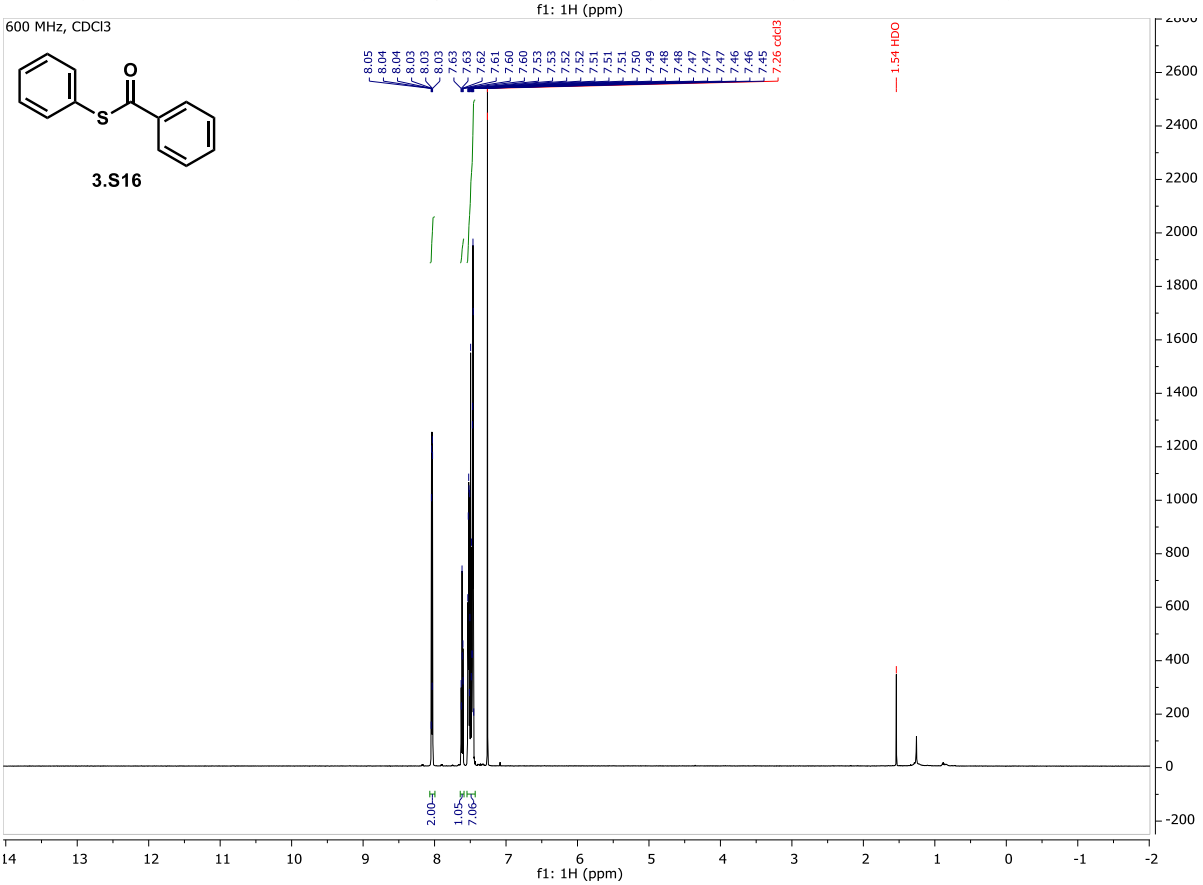

400 MHz, CDCl3
3.S11
f1: 1H (ppm)

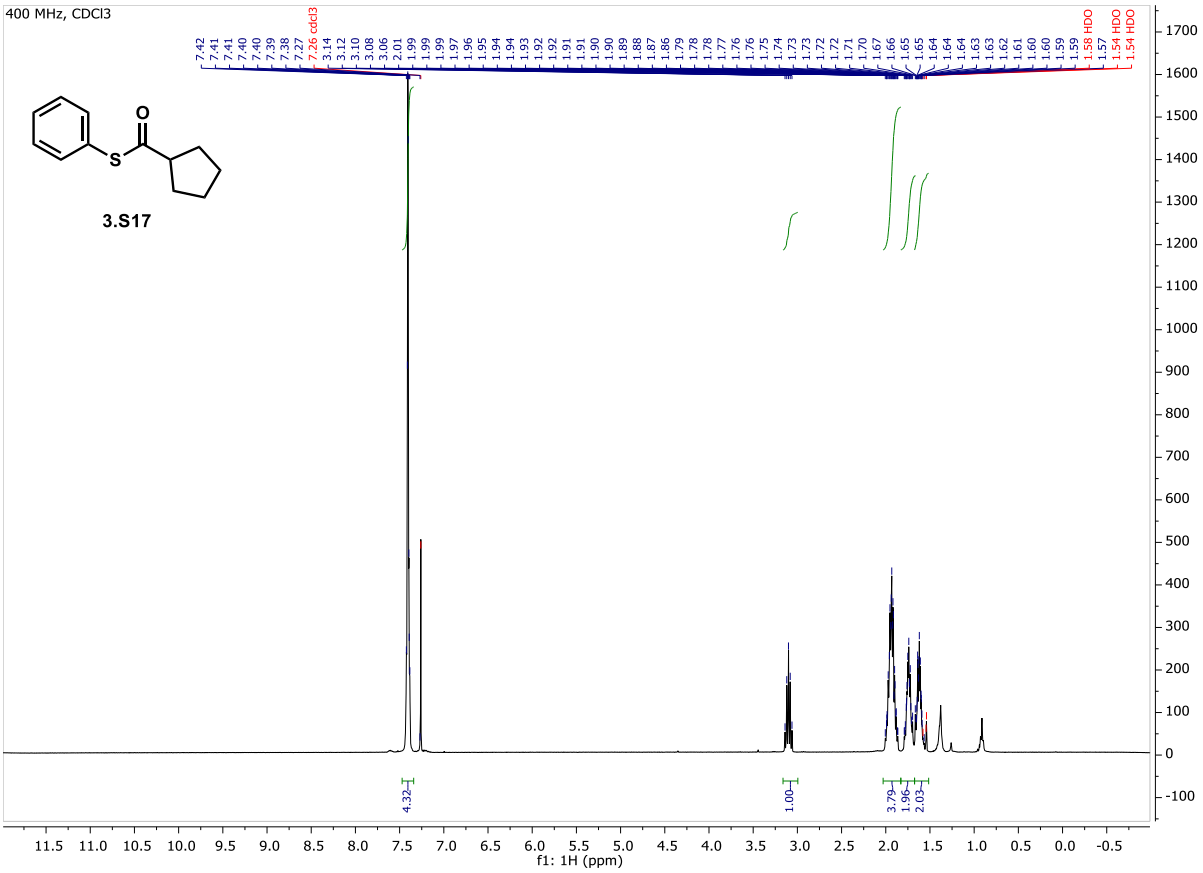

600 MHz, CDCl3
3.S12
f1: 1H (ppm)

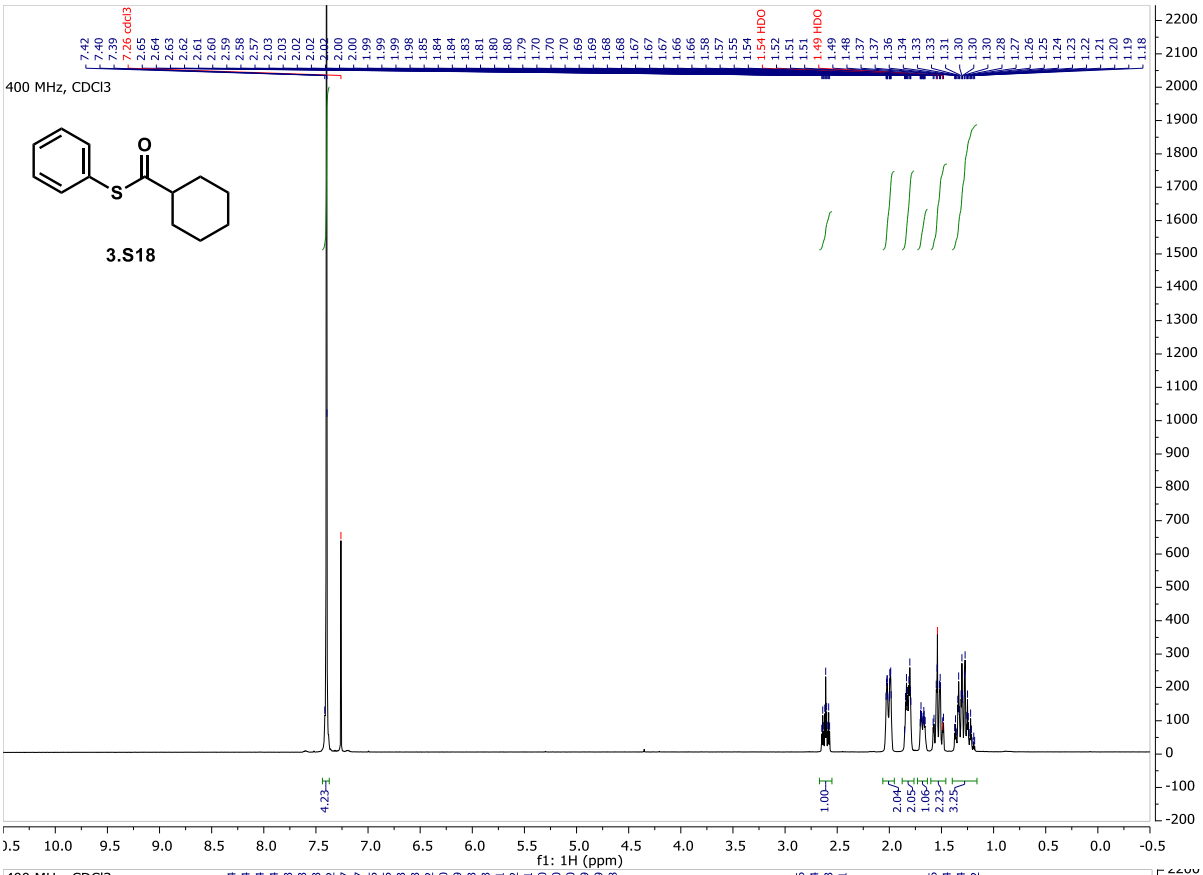

600 MHz, CDCl3
O
S
Me
3.S13
7.41
7.26 cdcl3
2.66
2.65
2.64
1.72
1.71
1.70
1.37
1.37
1.36
1.35
1.34
1.32
1.32
1.31
1.31
1.30
1.29
1.28
1.28
1.27
1.26
1.26
0.90
0.89
0.87
4.30
2.00
2.26
8.11
3.07
f1: 1H (ppm)

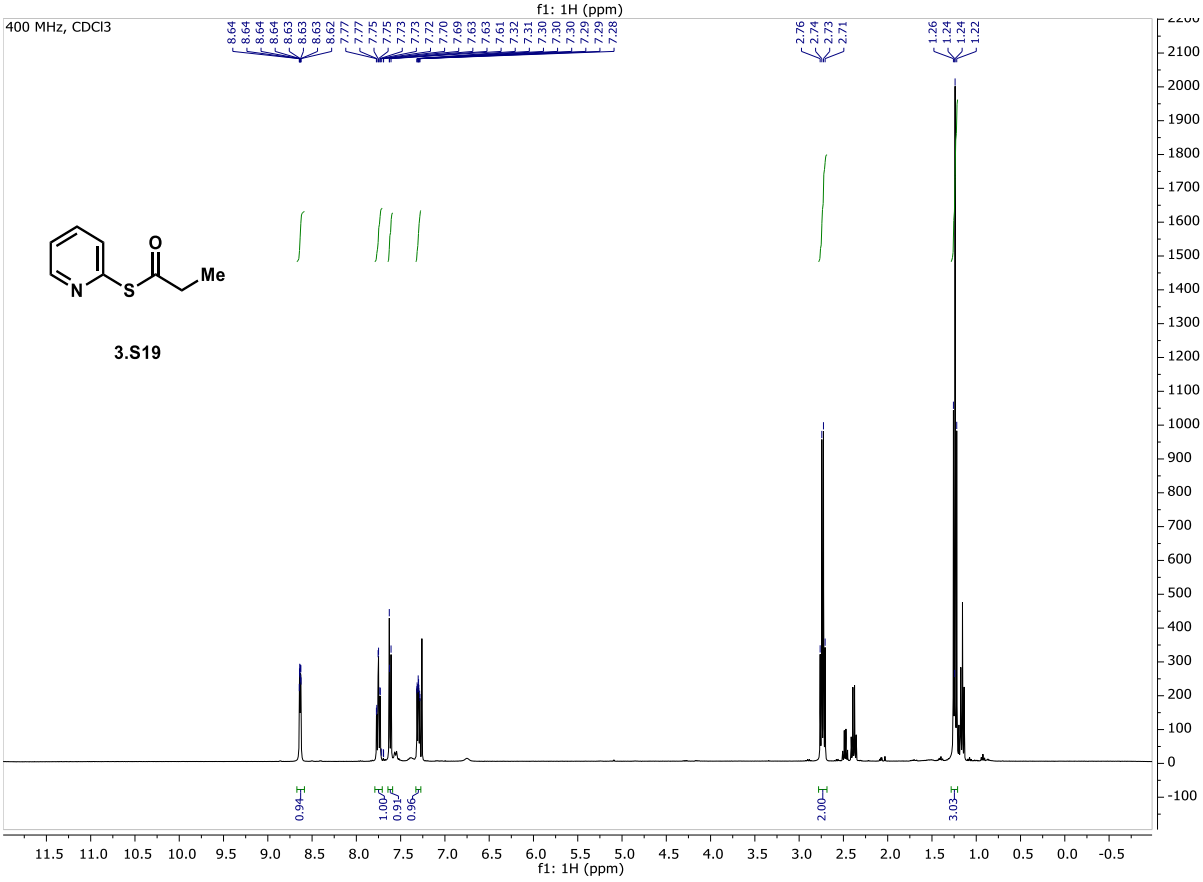

400 MHz, CDCl3
O
S
Me
Me
3.S14
7.41
7.26
2.92
2.90
2.89
2.87
2.85
2.83
2.82
1.28
1.26
4.59
0.78
6.00
f1: 1H (ppm)

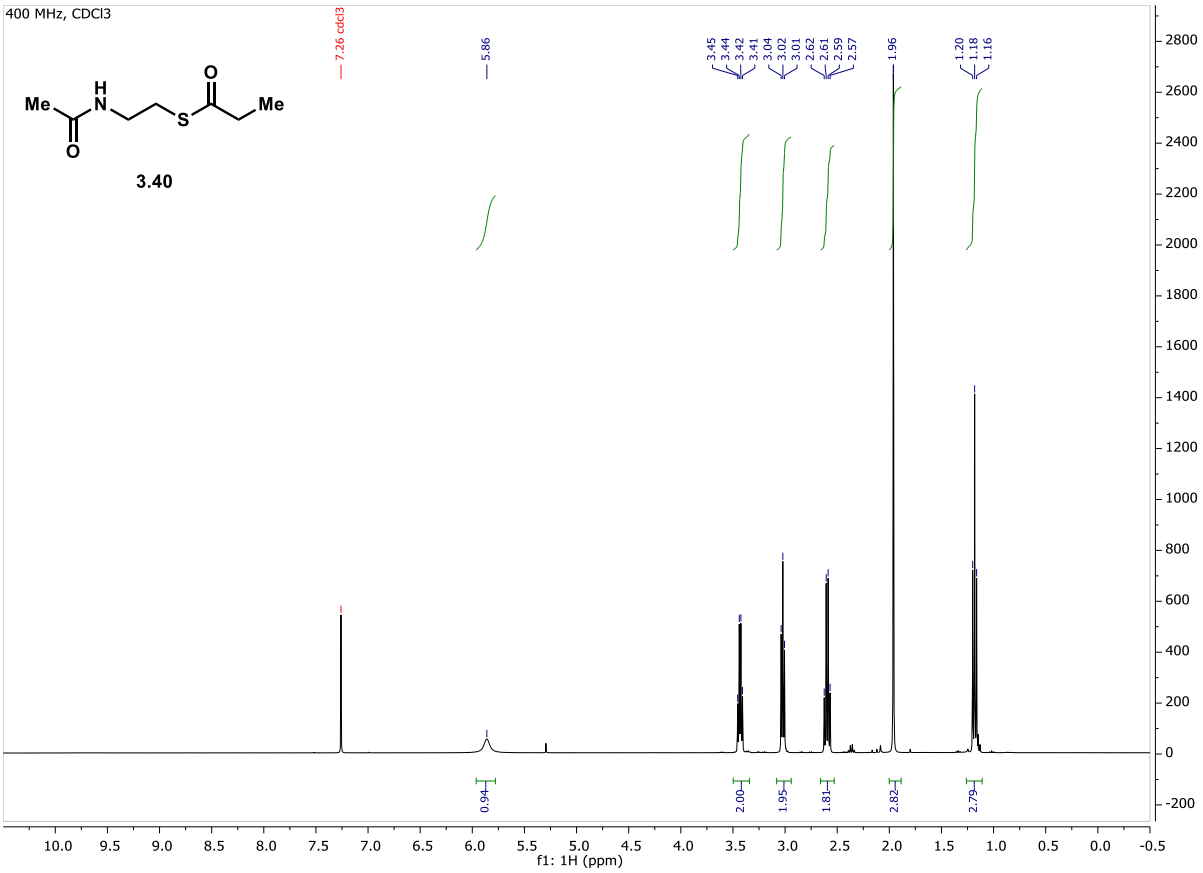

600 MHz, CDCl3
Me
O
S
Me
3.S15
7.41
7.41
7.26 cdcl3
2.55
2.53
2.25
2.24
2.23
2.22
2.21
2.19
2.18
1.54 HDO
1.01
1.00
4.12
1.91
1.00
5.60
f1: 1H (ppm)

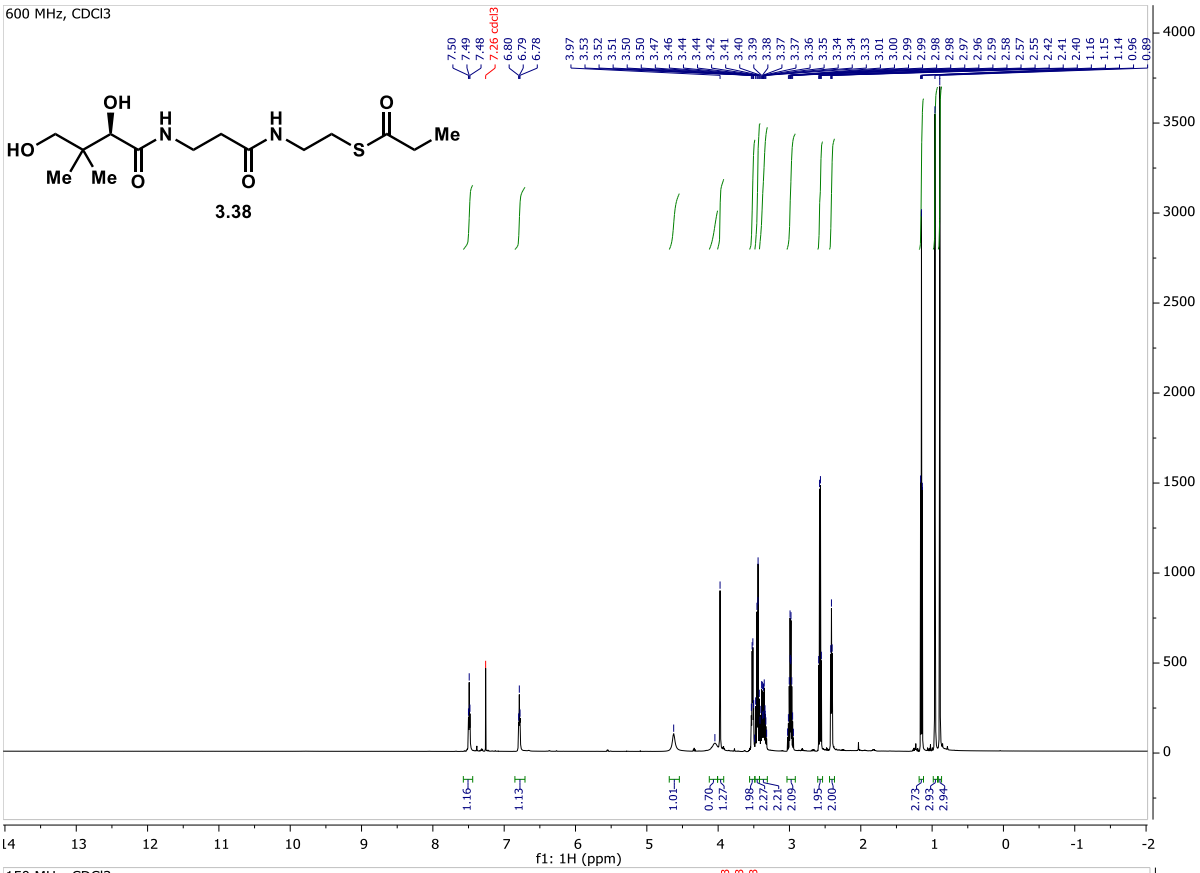

600 MHz, CDCl3
O
S
3.S16
8.05
8.04
8.04
8.03
8.03
8.03
7.63
7.63
7.62
7.61
7.60
7.60
7.53
7.53
7.52
7.52
7.51
7.51
7.51
7.50
7.49
7.48
7.48
7.47
7.47
7.46
7.46
7.45
7.26 cdcl3
1.54 HDO
2.00
1.05
7.06
f1: 1H (ppm)

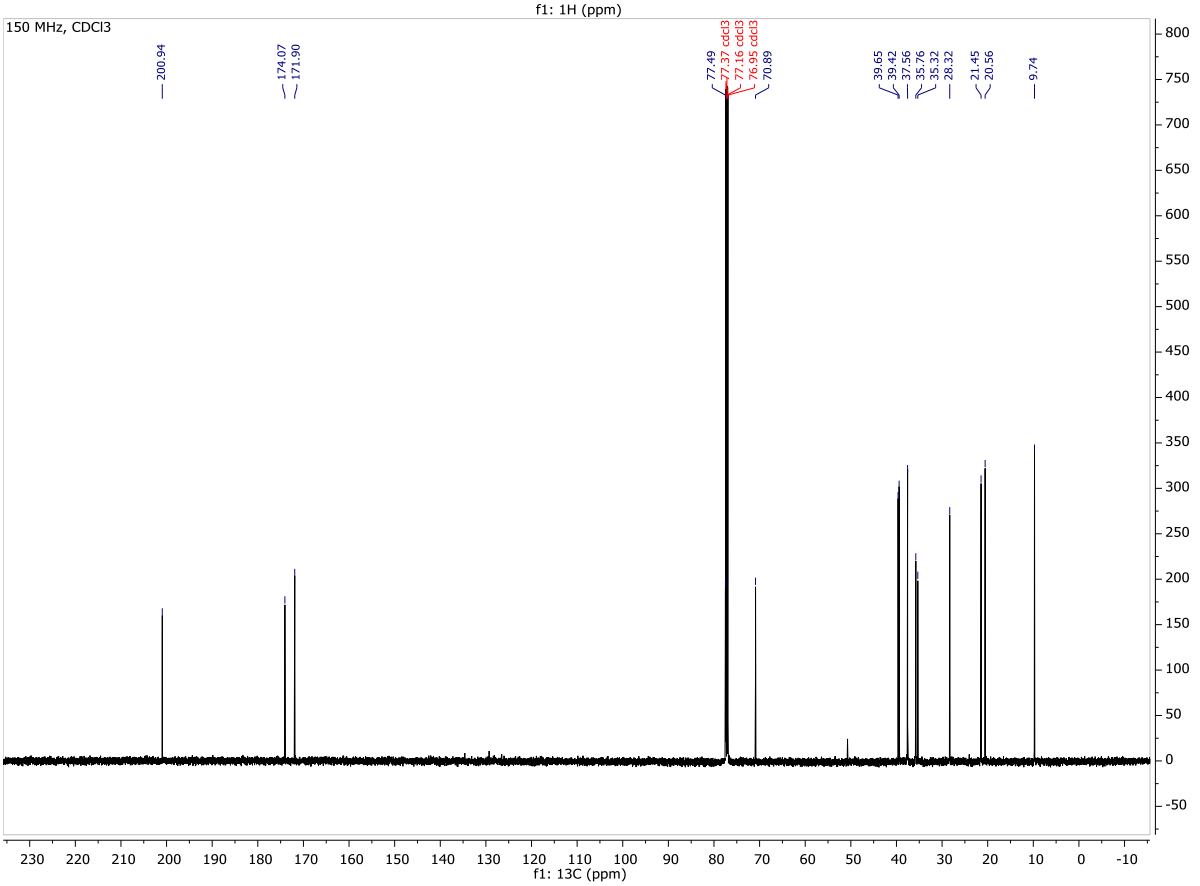

400 MHz, CDCl3
3.S17
f1: 1H (ppm)

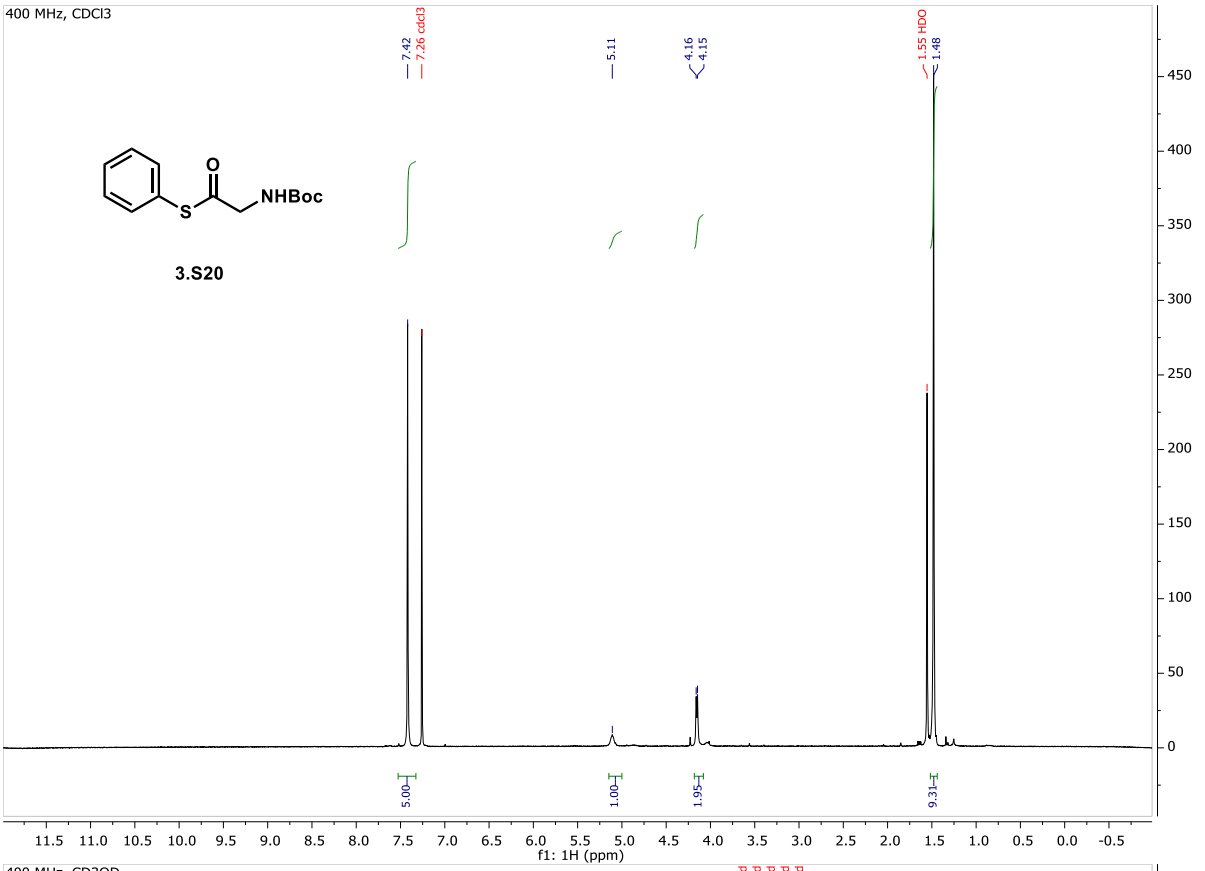

400 MHz, CDCl3
3.S18
f1: 1H (ppm)

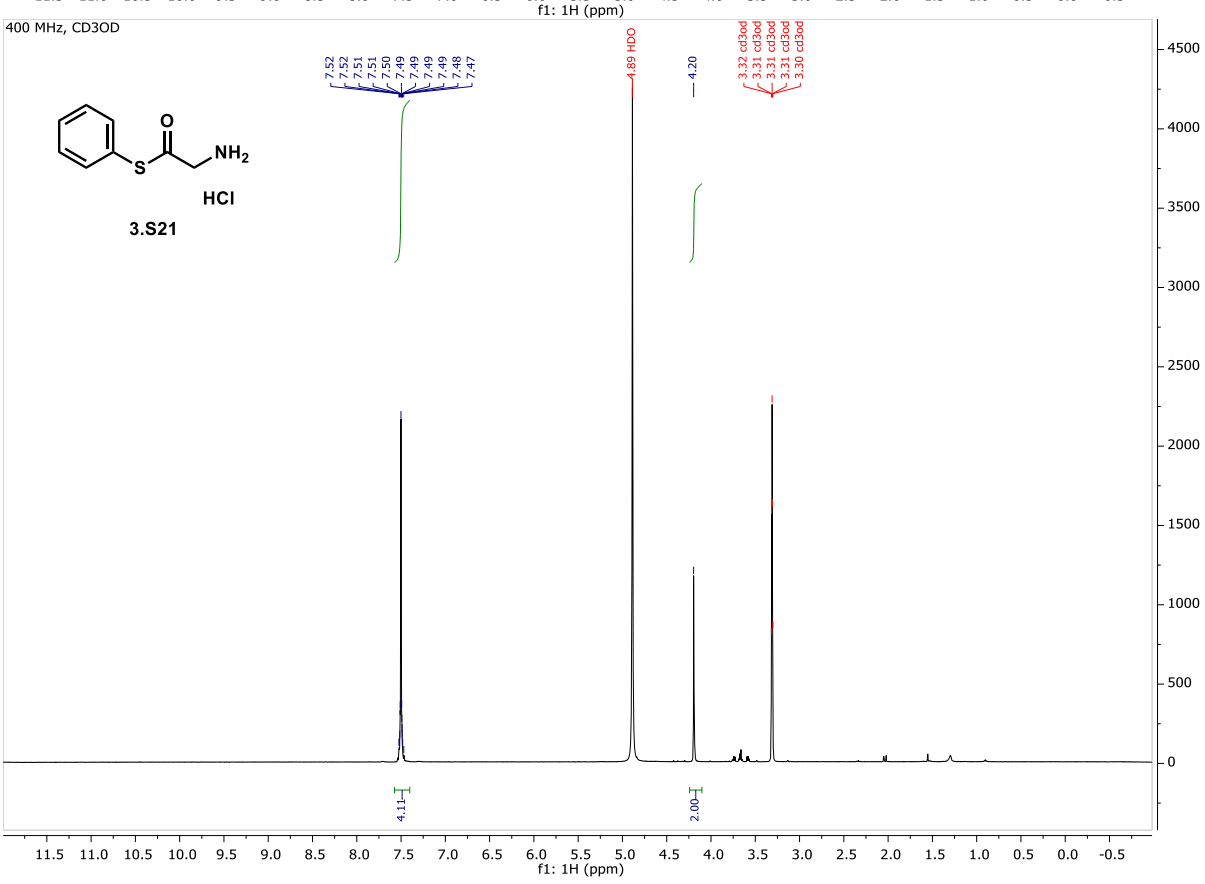

400 MHz, CDCl3
3.S19
f1: 1H (ppm)

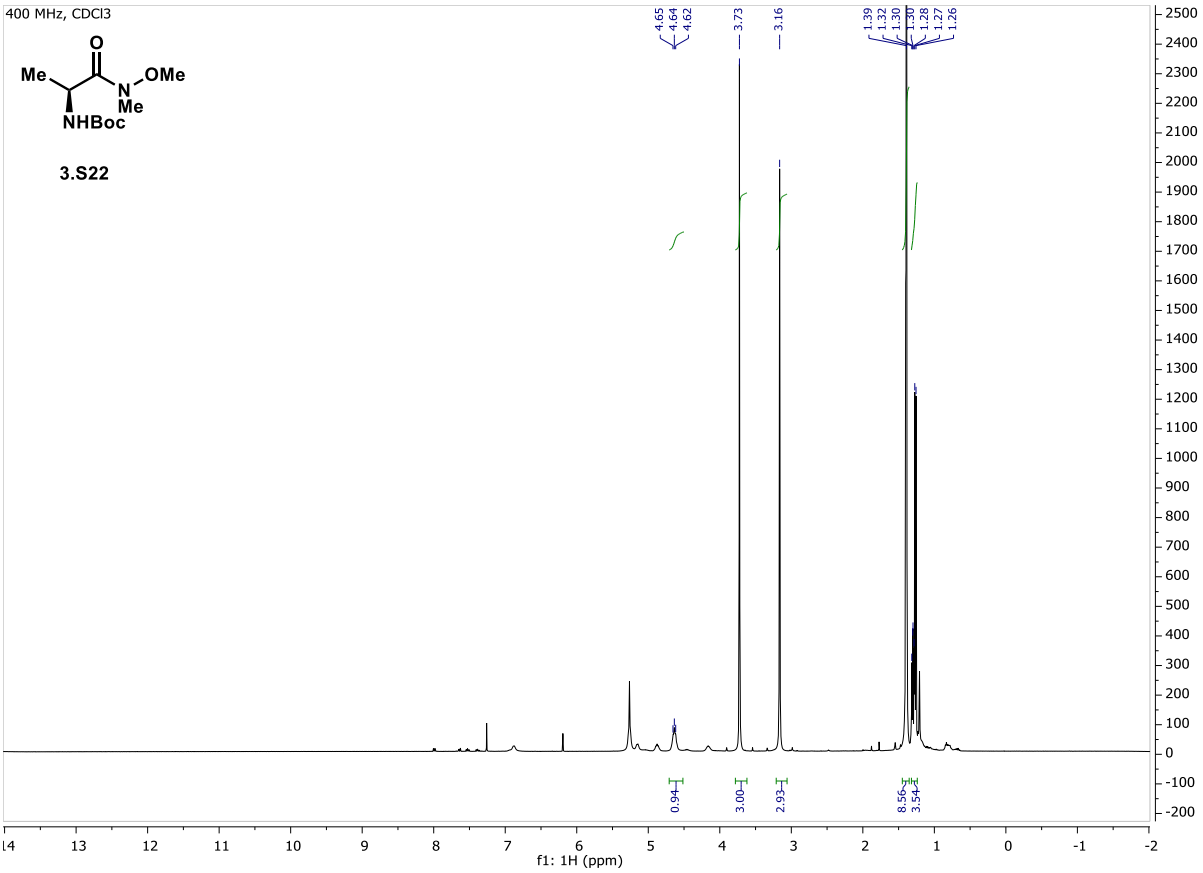

400 MHz, CDCl3
7.26 cdcl3
5.86
3.45
3.44
3.42
3.41
3.04
3.02
3.01
2.62
2.61
2.59
2.57
1.96
1.20
1.18
1.16
Me
H
N
O
S
O
Me
O
3.40
0.94
2.00
1.95
1.81
2.82
2.79
f1: 1H (ppm)

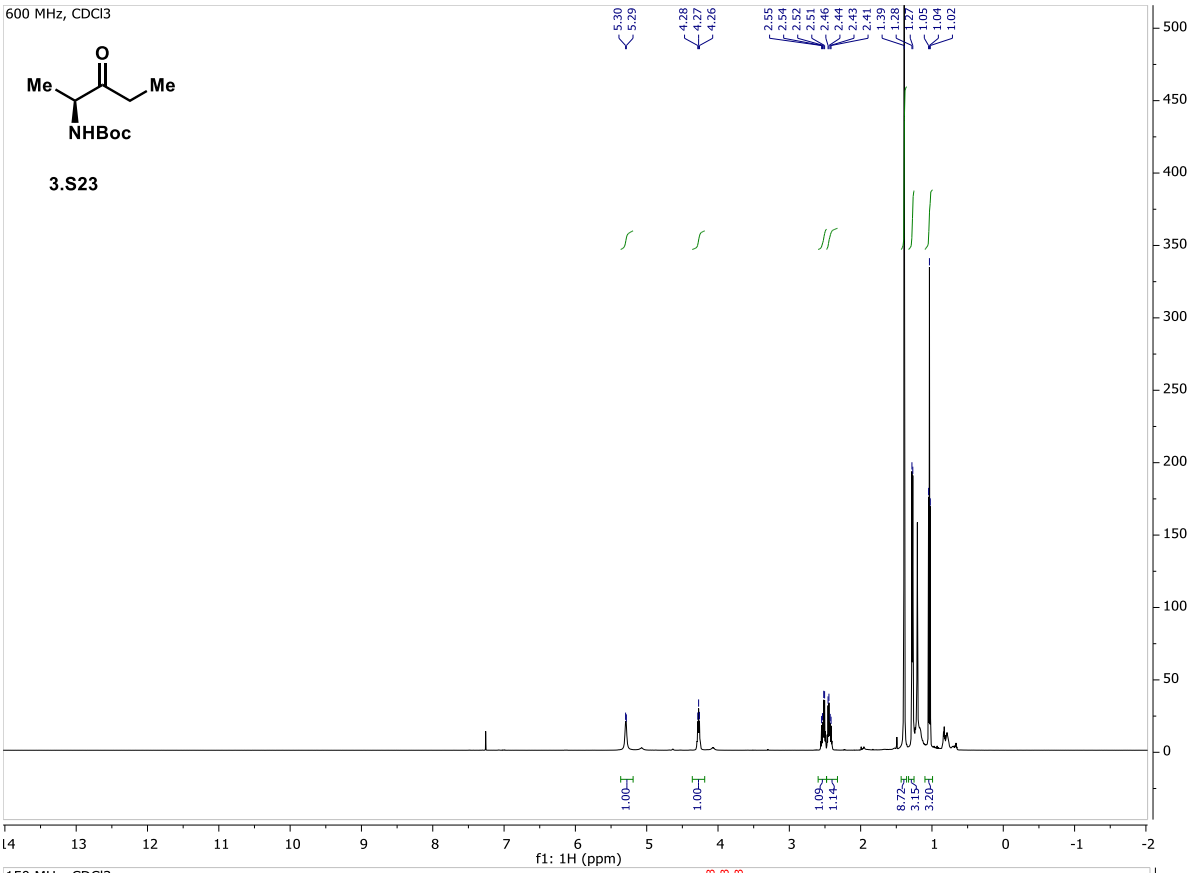

600 MHz, CDCl3
3.38
f1: 1H (ppm)

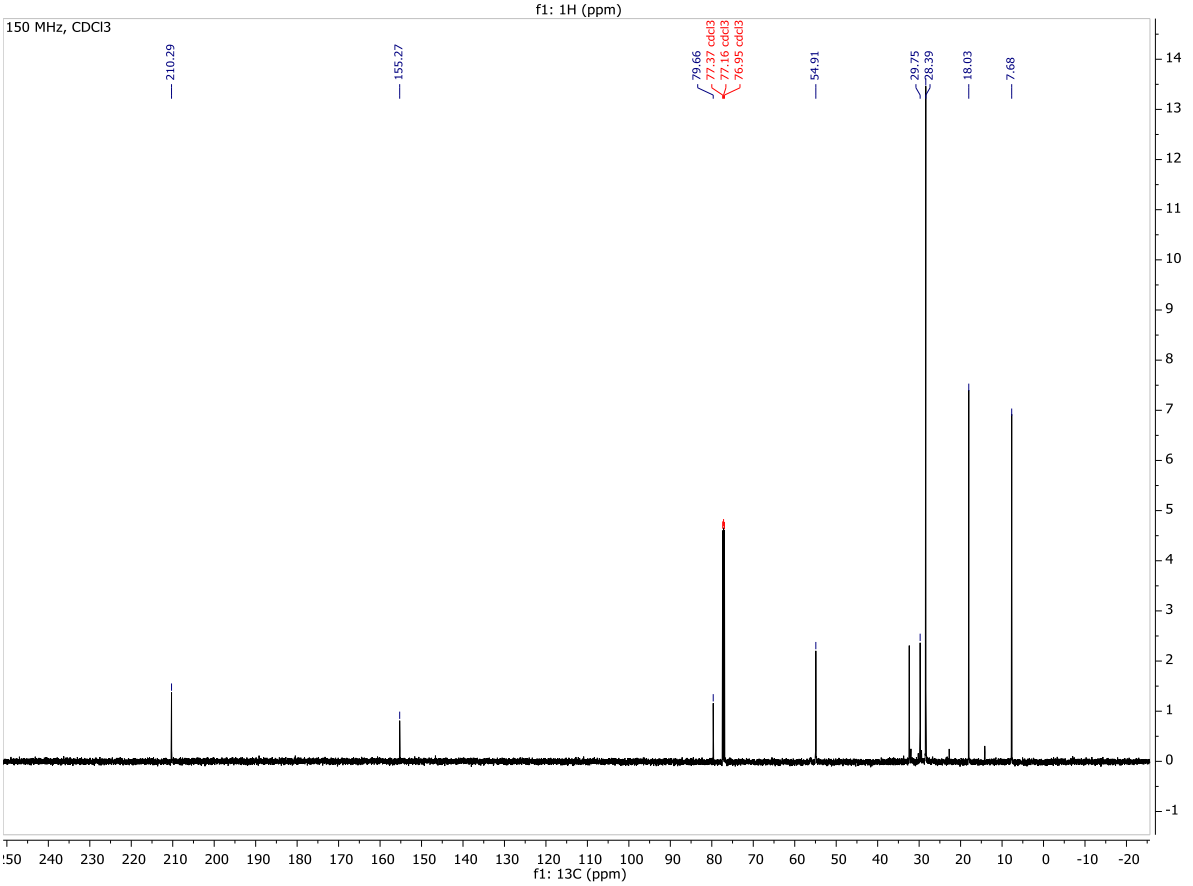

150 MHz, CDCl3
f1: 13C (ppm)

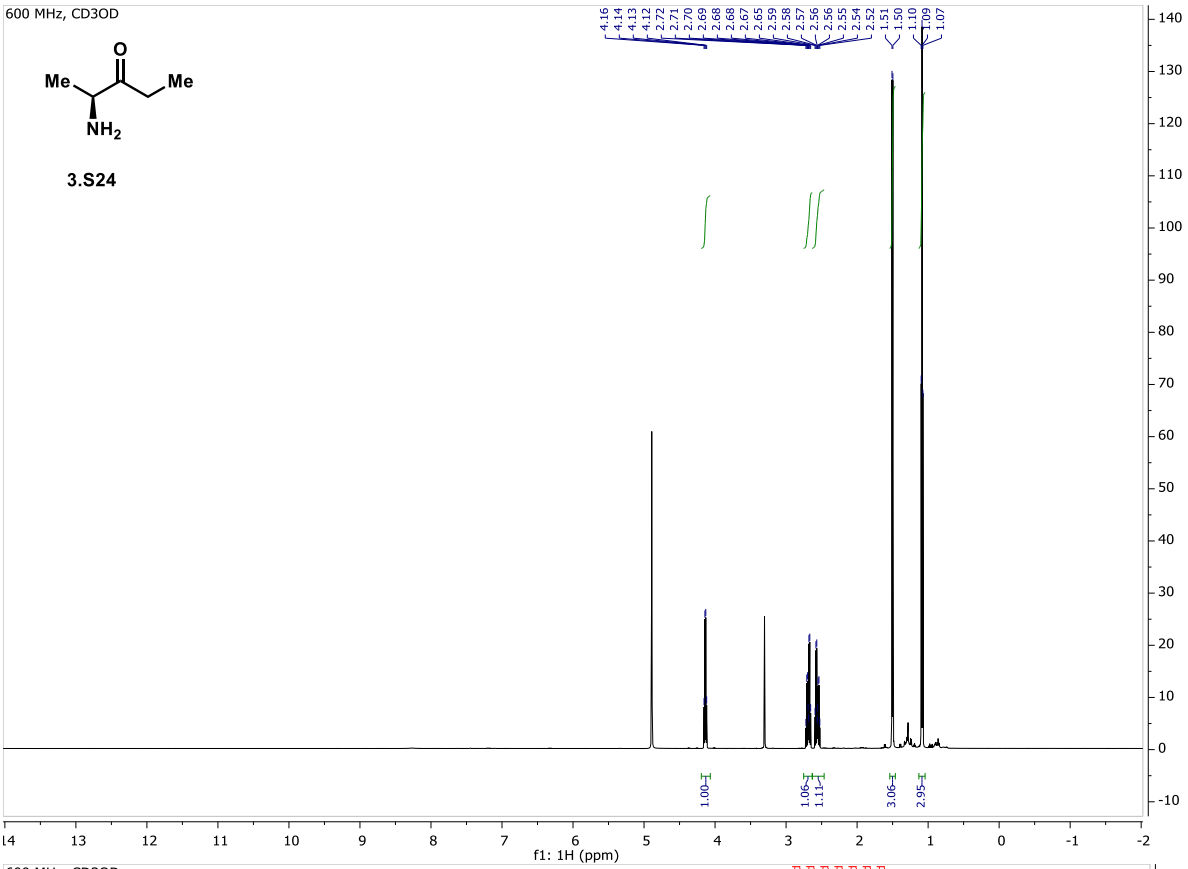

400 MHz, CDCl3
7.42
7.26 cdcl3
5.11
4.16
4.15
1.55 HDO
1.48
O
S
NHBoc
3.S20
5.00
1.00
1.95
9.31
f1: 1H (ppm)

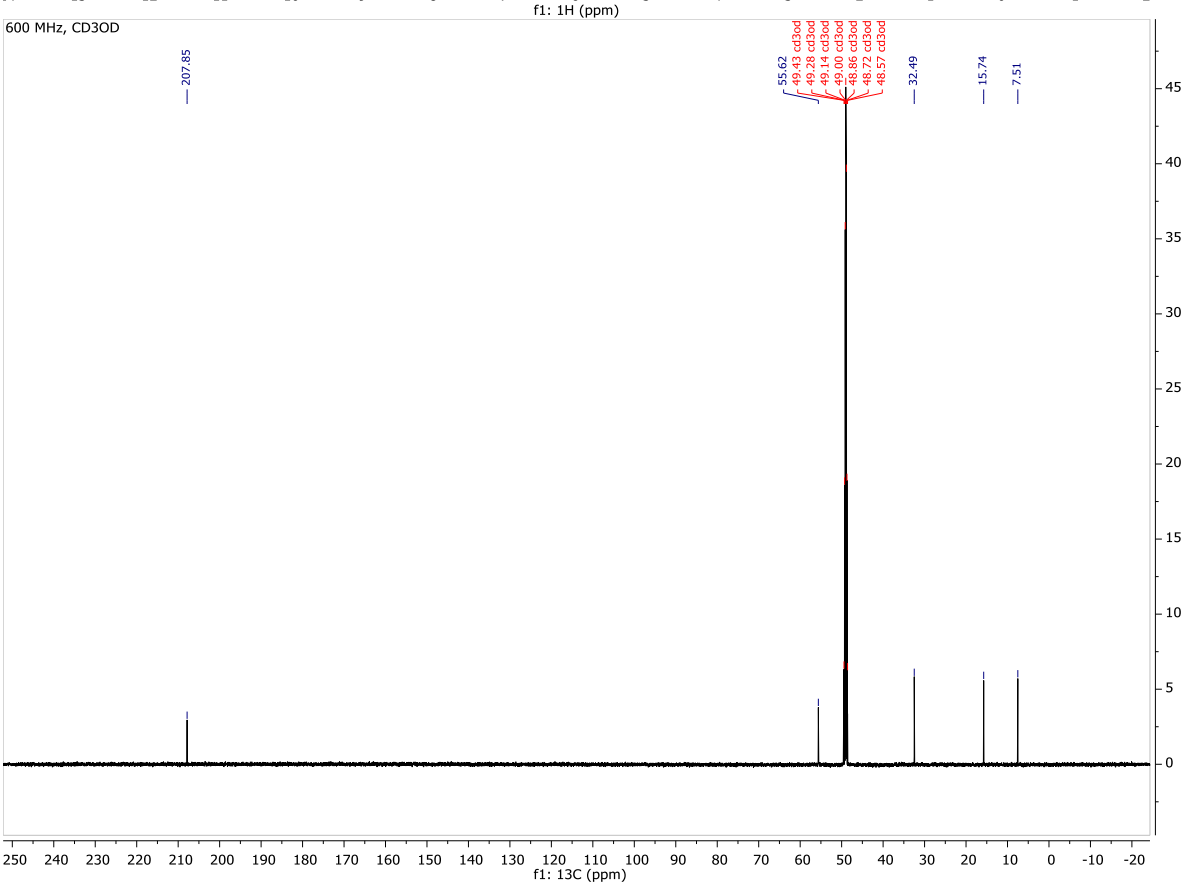

400 MHz, CD3OD
7.52
7.52
7.51
7.51
7.50
7.49
7.49
7.49
7.49
7.48
7.47
4.89 HDO
4.20
3.32 cd3od
3.31 cd3od
3.31 cd3od
3.31 cd3od
3.30 cd3od
O
S
NH2
HCl
3.S21
4.11
2.00
f1: 1H (ppm)

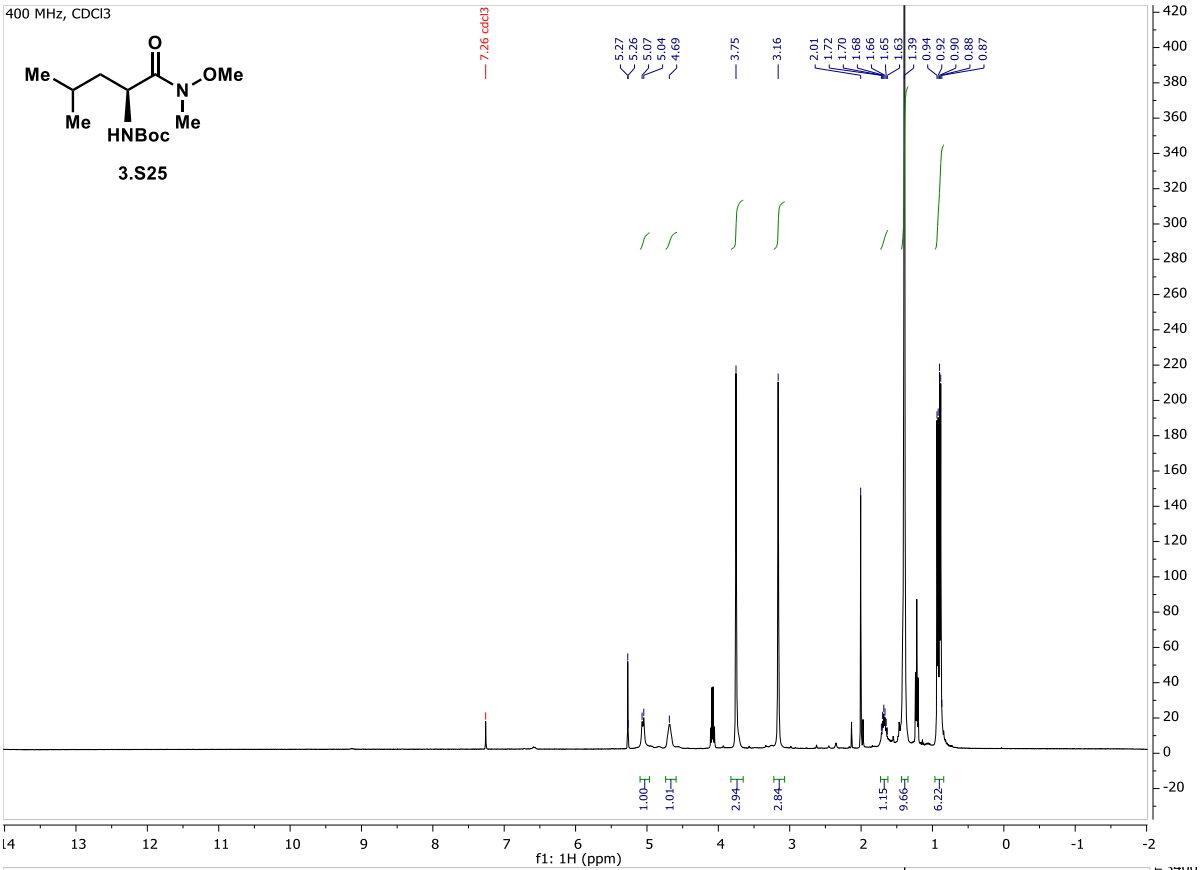

400 MHz, CDCl3
Me
O
N
OMe
Me
NHBoc
3.S22
4.65
4.64
4.62
3.73
3.16
1.39
1.32
1.30
1.30
1.28
1.27
1.26
0.94
3.00
2.93
8.56
3.54
f1: 1H (ppm)

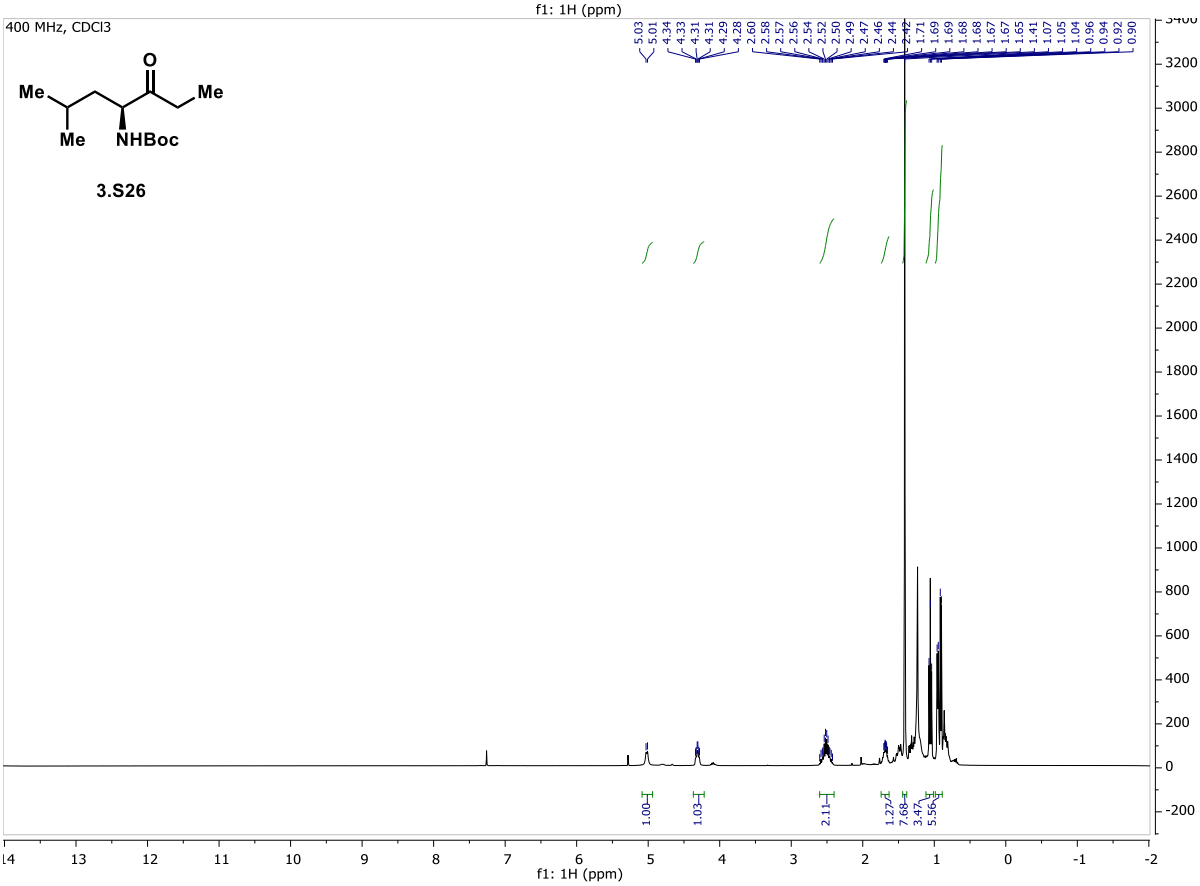

600 MHz, CDCl3
Me
Me
O
NHBoc
3.S23
5.30
5.29
4.28
4.27
4.26
2.55
2.54
2.52
2.51
2.46
2.44
2.43
2.41
1.39
1.28
1.27
1.05
1.04
1.02
1.00
1.00
1.09
1.14
8.72
3.15
3.20
f1: 1H (ppm)

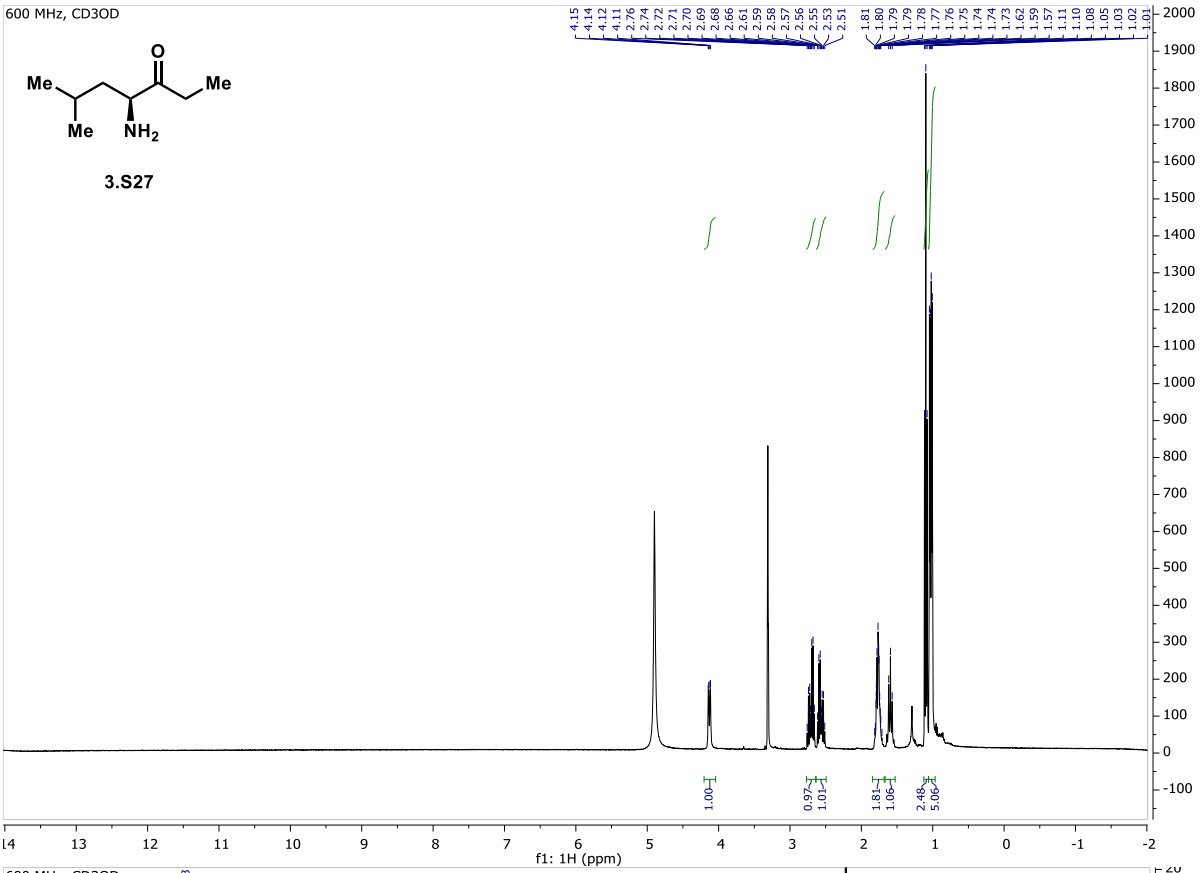

150 MHz, CDCl3
210.29
155.27
79.66
77.37 cdcl3
77.16 cdcl3
76.95 cdcl3
54.91
29.75
28.39
18.03
7.68
f1: 13C (ppm)

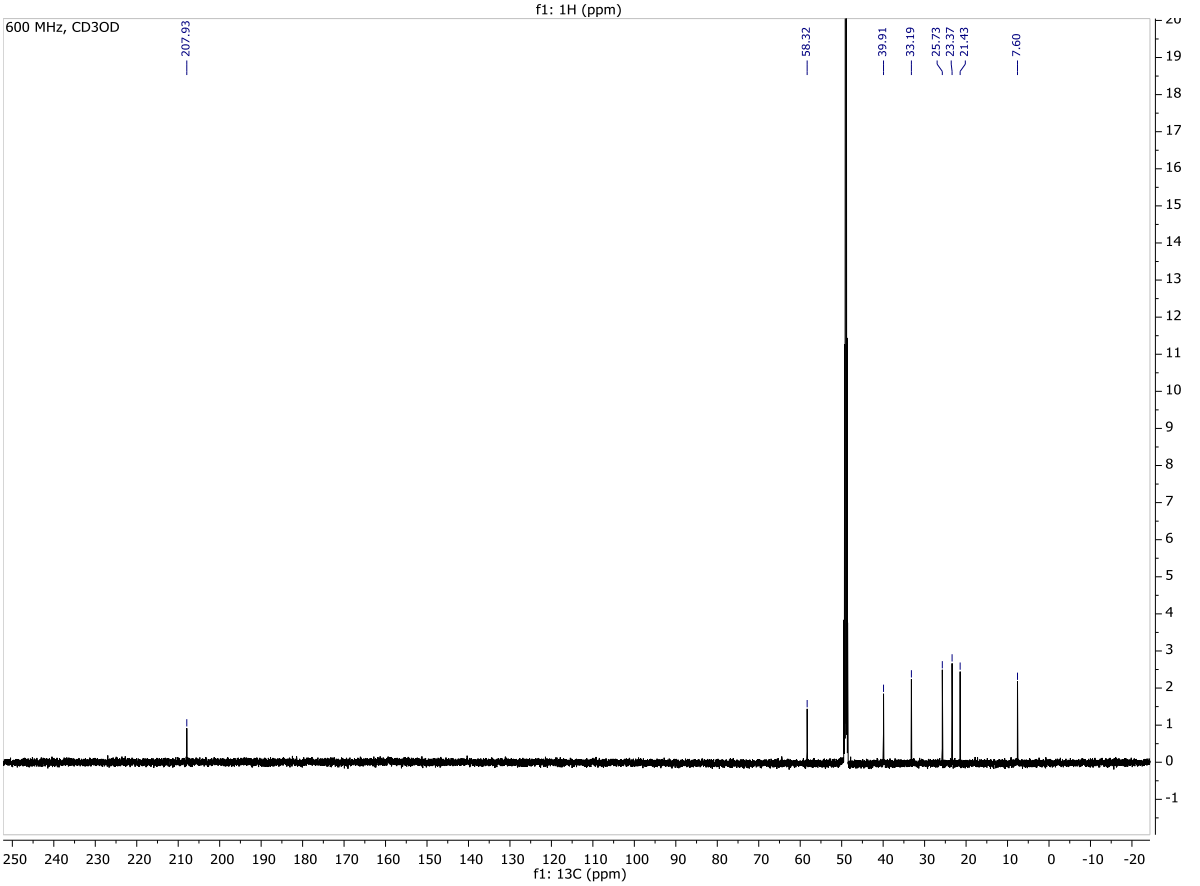

600 MHz, CD3OD
Me
Me
O
NH2
3.S24
4.16
4.14
4.13
4.12
2.72
2.71
2.70
2.69
2.68
2.68
2.67
2.65
2.59
2.58
2.57
2.56
2.56
2.55
2.54
2.52
1.51
1.50
1.10
1.09
1.07
1.00
1.06
1.11
3.06
2.95
f1: 1H (ppm)

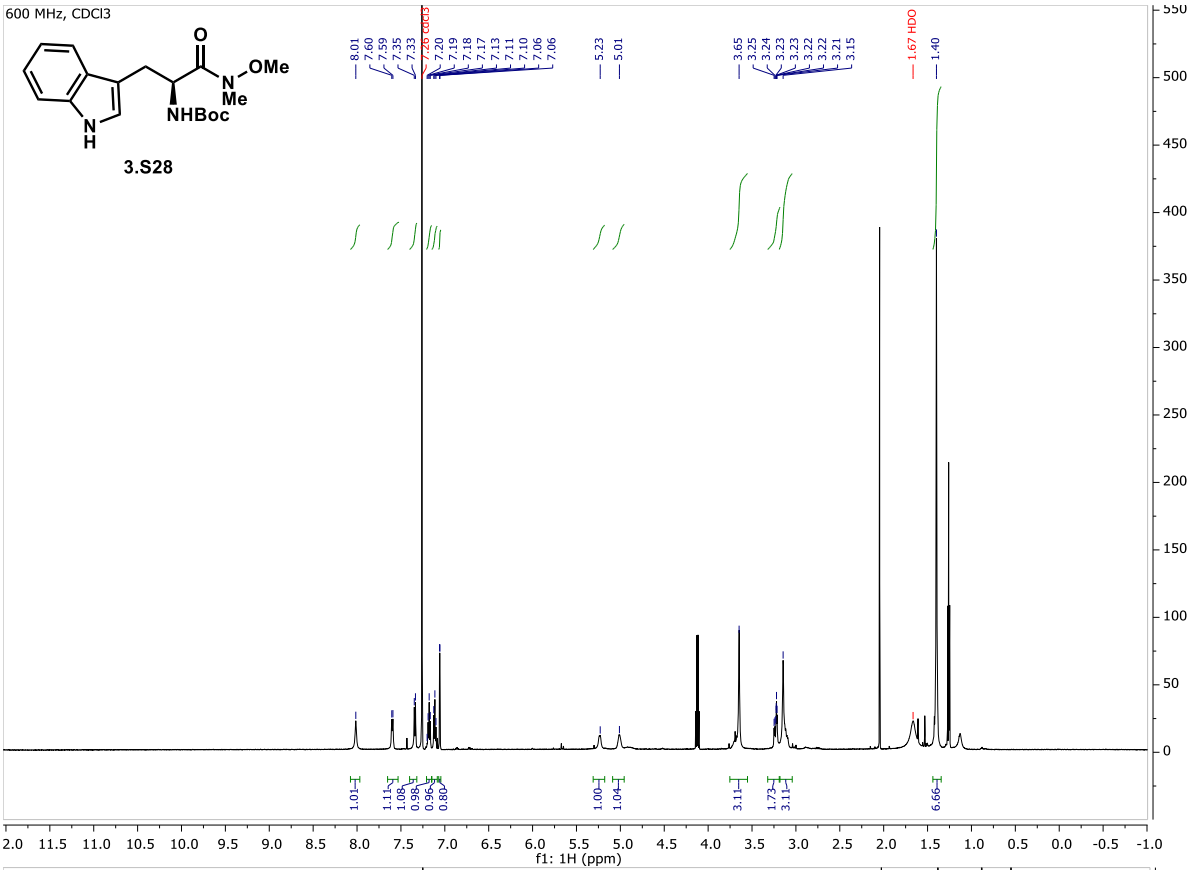

600 MHz, CD3OD
207.85
55.62
49.43 cd3od
49.28 cd3od
49.14 cd3od
49.00 cd3od
48.86 cd3od
48.72 cd3od
48.57 cd3od
32.49
15.74
7.51
f1: 13C (ppm)

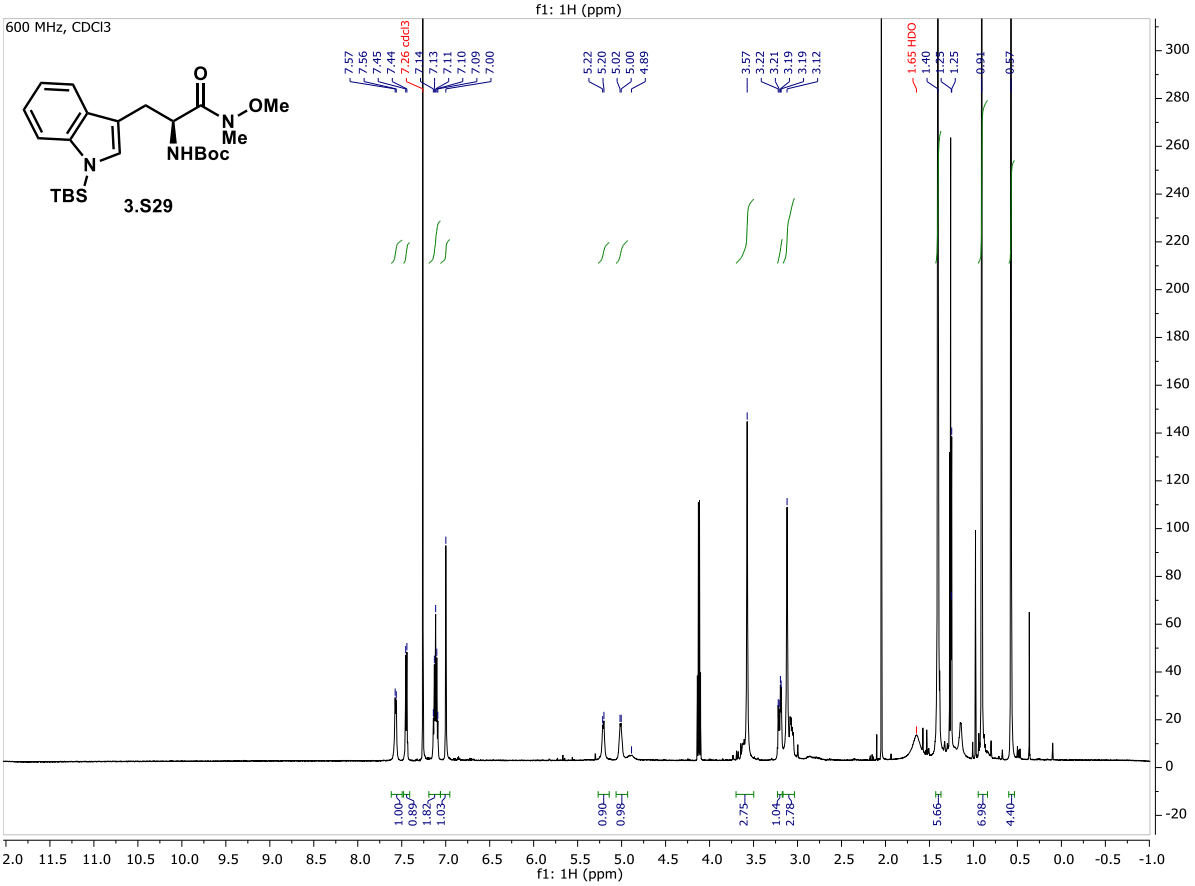

400 MHz, CDCl3
Me
Me
HNBoc
O
N
OMe
Me
3.S25
7.26 cdcl3
5.27
5.26
5.07
5.04
4.69
3.75
3.16
2.01
1.72
1.70
1.68
1.66
1.65
1.63
1.39
0.94
0.92
0.90
0.88
0.87
1.00
1.01
2.94
2.84
1.15
9.66
6.22
f1: 1H (ppm)

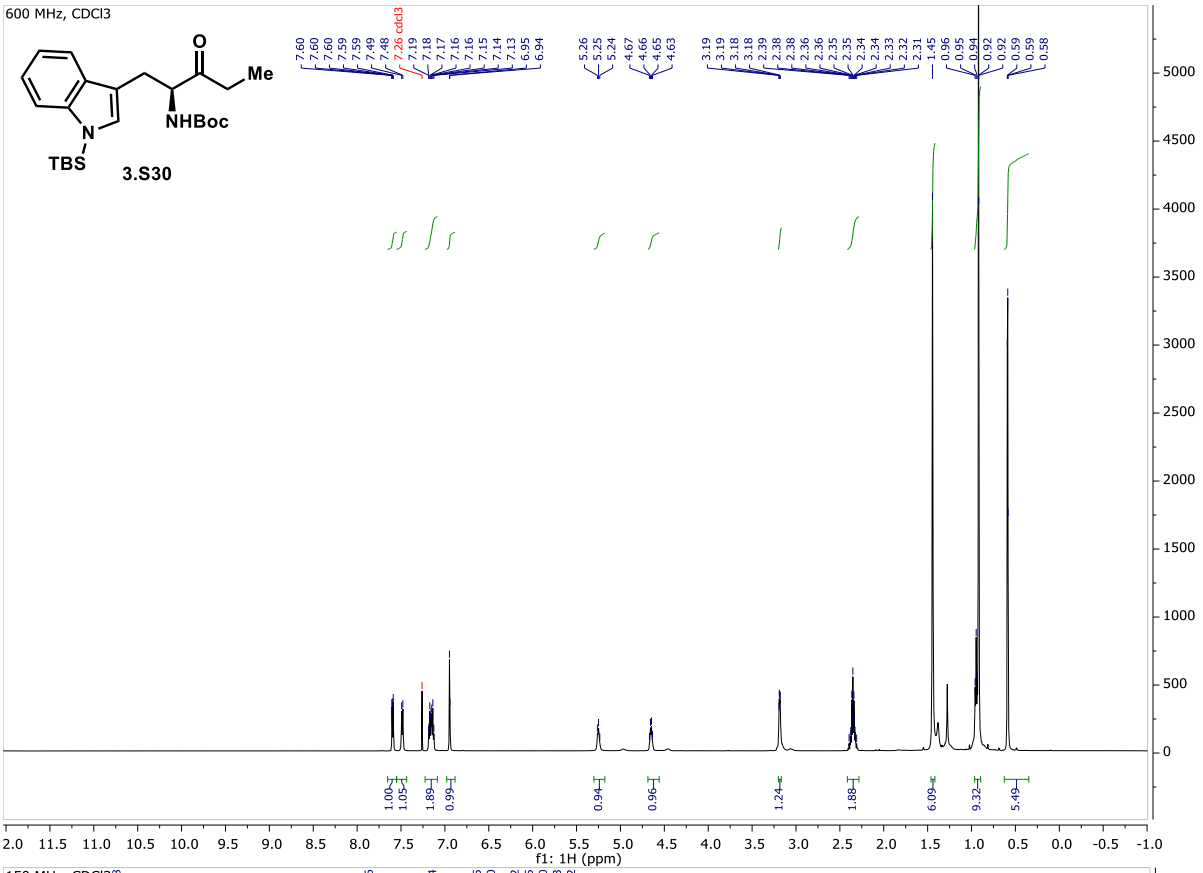

400 MHz, CDCl3
Me
Me
NHBoc
O
Me
3.S26
5.03
5.01
4.34
4.33
4.31
4.31
4.29
4.28
2.60
2.58
2.57
2.56
2.54
2.52
2.50
2.49
2.47
2.46
2.44
2.42
1.71
1.69
1.69
1.68
1.68
1.67
1.67
1.65
1.41
1.07
1.05
1.04
0.96
0.94
0.92
0.90
1.00
1.03
2.11
1.27
7.68
3.47
5.56
f1: 1H (ppm)

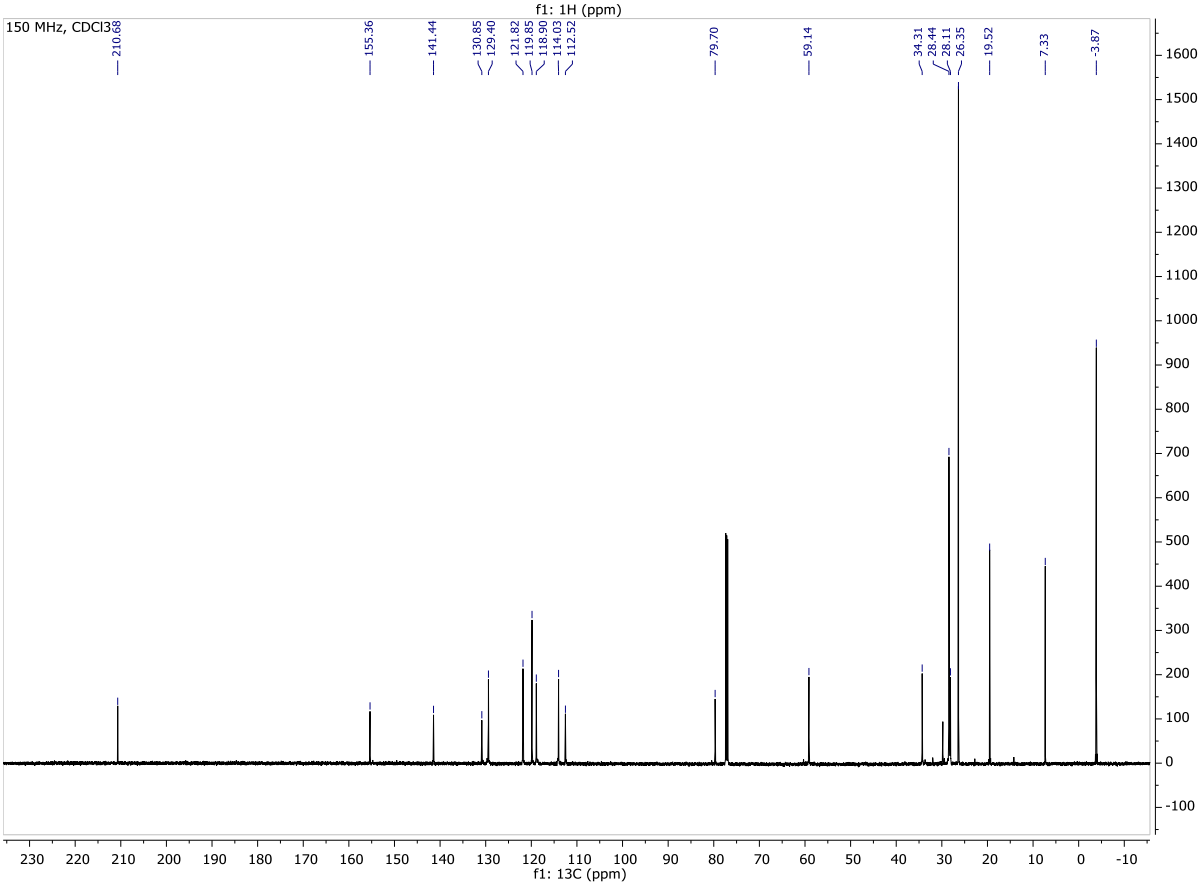

600 MHz, CD3OD
Me
Me
Me
O
NH2
3.S27
f1: 1H (ppm)

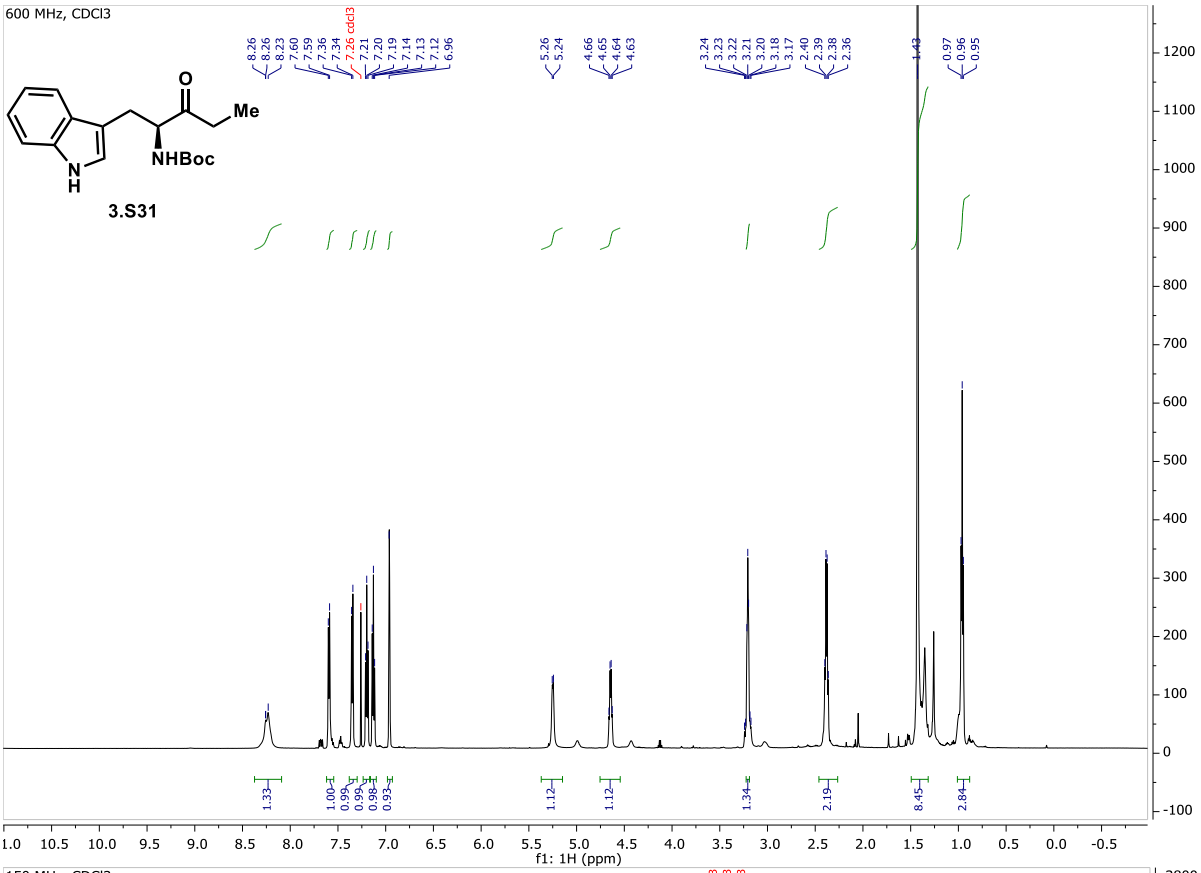

600 MHz, CD3OD
207.93
58.32
39.91
33.19
25.73
23.37
21.43
7.60
f1: 13C (ppm)

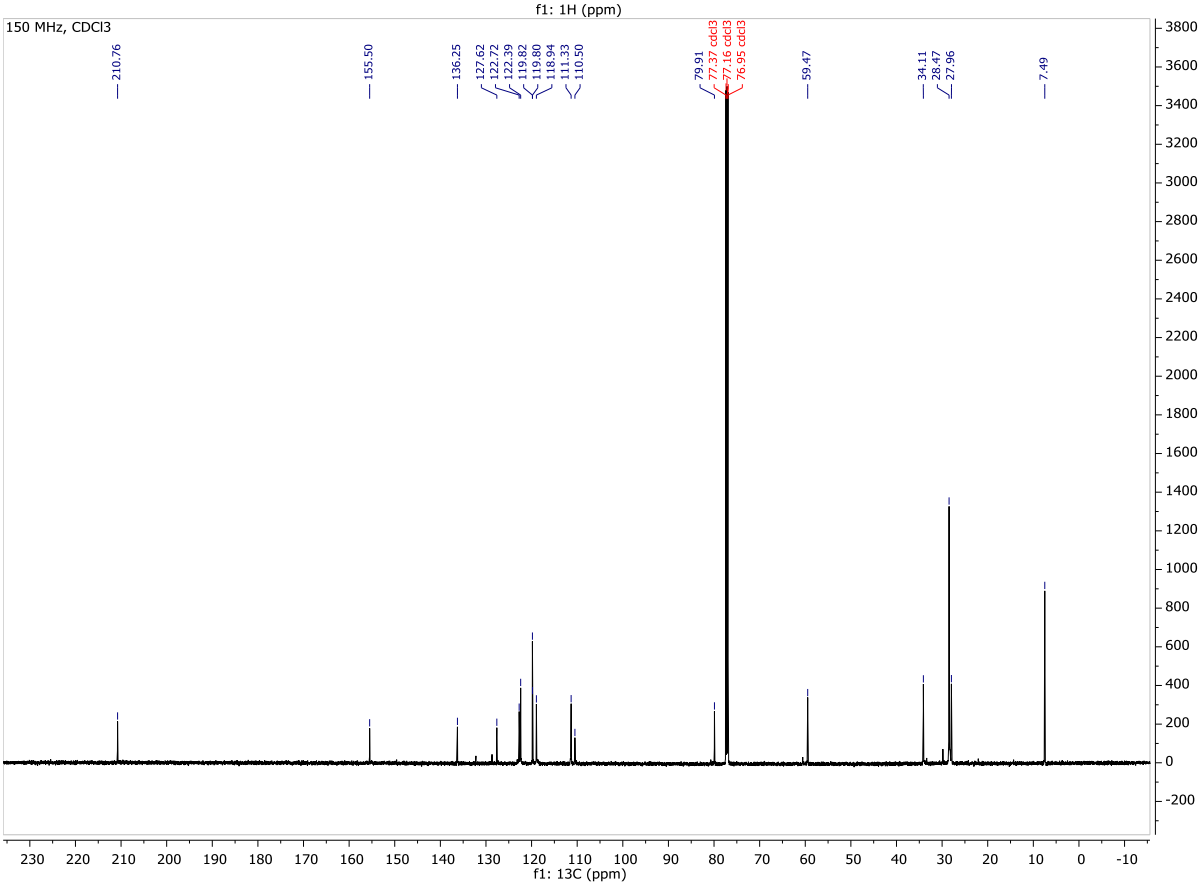

600 MHz, CDCl3
3.S28
f1: 1H (ppm)

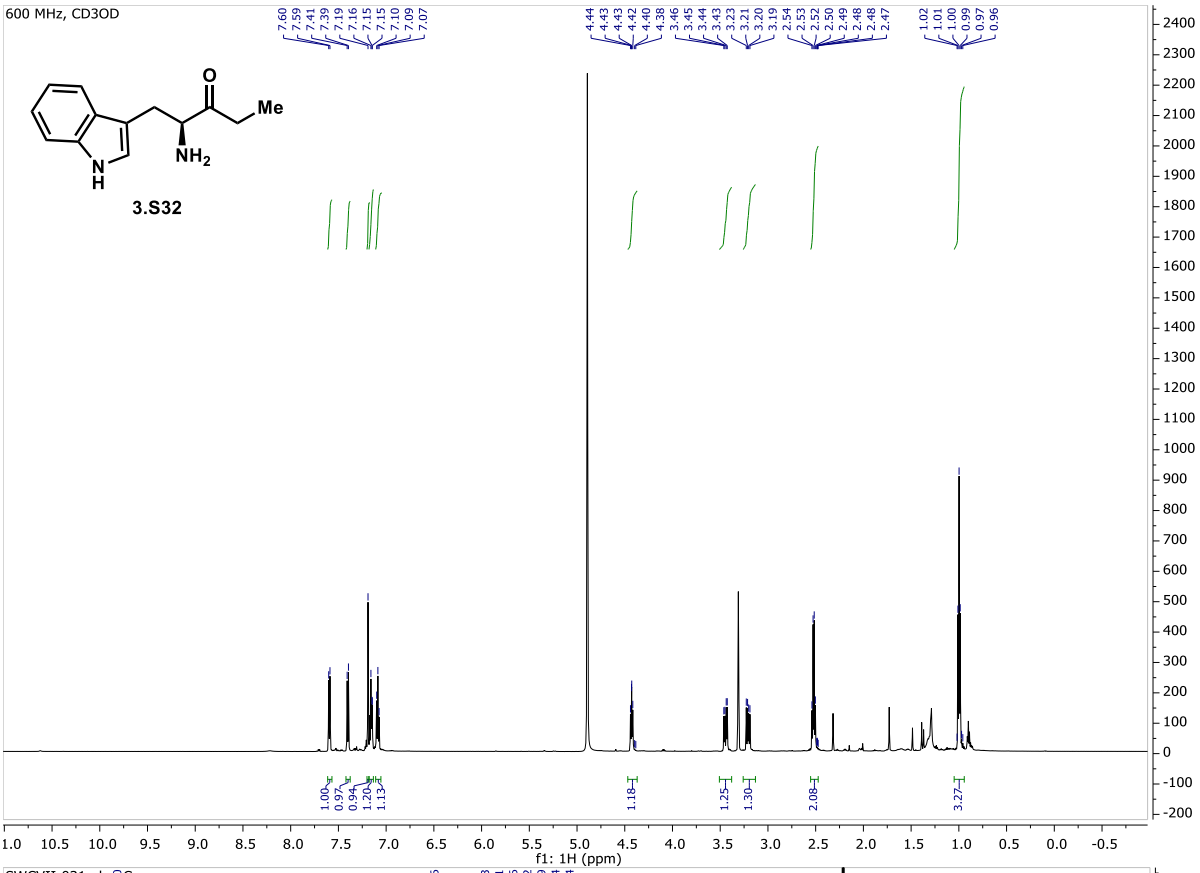

600 MHz, CDCl3
TBS
3.S29
f1: 1H (ppm)

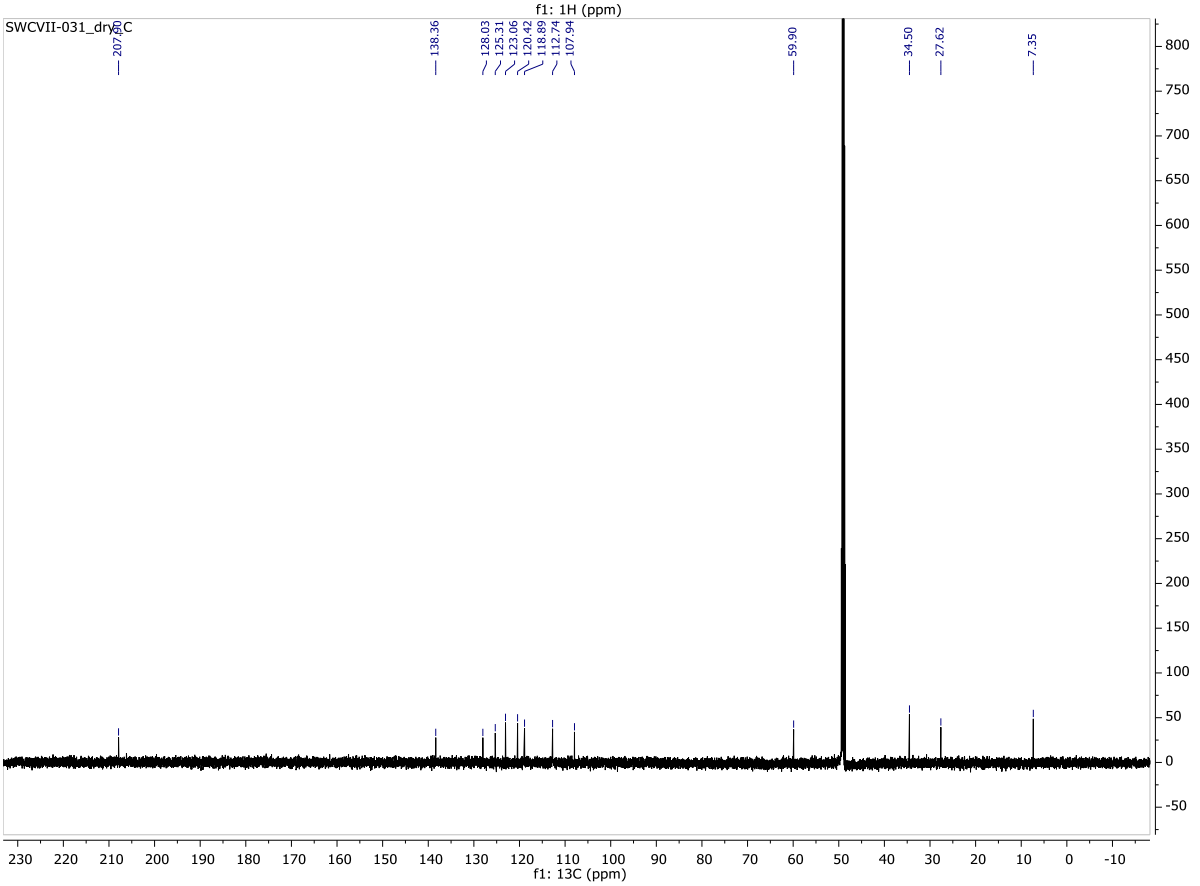

600 MHz, CDCl3
O
Me
NHBoc
N
TBS
3.S30
f1: 1H (ppm)


150 MHz, CDCl3
f1: 13C (ppm)


600 MHz, CDCl3
Me
NHBoc
N
H
3.S31
f1: 1H (ppm)


150 MHz, CDCl3
f1: 13C (ppm)


600 MHz, CD3OD
3.S32
f1: 1H (ppm)


SWCVII-031_dry_C
f1: 13C (ppm)

II. Cloning, protein expression and purification

A. General information

Escherichia coli cloning strains DH10B or DH5α (Invitrogen) were used for DNA propagation. Proteins were expressed in *E. coli* strains BL21(DE3) pRARE (customized for streptomycin and spectinomycin resistance by the University of Michigan Center for Structural Biology), BAP1[87] or BAP1 previously transformed with pGro7 (Takara Bio) after transformation with genes subcloned into pET28b (Novagen) or pMCSG7 (provided by the University of Michigan Center for Structural Biology) plasmids. Cells were grown in either Luria-Bertani broth (LB) or Terrific broth (TB) supplemented with the corresponding antibiotics kanamycin (50 μg/mL), ampicillin (100 μg/mL), chloramphenicol (35 μg/mL) or spectinomycin (80 μg/mL) purchased from Gold Biotechnology.

All primers were purchased from Integrated DNA Technologies. Phusion HF polymerase and DpnI restriction enzyme were purchased from New England BioLabs. Ligation-independent cloning (LIC)-qualified T4 DNA polymerase was purchased from EMD Millipore. QIAquick PCR purification, gel extraction and miniprep kits were purchased from Qiagen. HisPur nickel-nitrilotriacetic acid (Ni-NTA) resin was purchased from Thermo Scientific. Proteins were concentrated using Amicon centrifugal filters purchased from EMD Millipore at 4,000 x *g*, 4 °C. PD-10 desalting columns were purchased from GE Healthcare. Protein samples were analyzed on Mini-PROTEAN TGX Gels (4-15%) from BioRad and visualized with Protein Ark Quick Coomassie Stain from Anatrace. Proteins were quantified with the Pierce 660 nm Assay Reagent from Thermo Scientific.

Purified maltose binding protein fusions of CoA biosynthetic proteins CoaA, CoaD and CoaE were kindly donated by Dr. Gregory Dodge of Prof. Janet Smith's group at the University of Michigan.

B. Sequence information

pET28b-s*xtA* (*Cylindrospermopsis raciborskii* T3) was obtained from the laboratory of Prof. Brett Neilan (University of Newcastle).

pET28b-s*xtA* ACP-AONS (*Aphanizomenon gracile*), pET28b-s*xtA* ACP-AONS (*Aphanizomenon gracile*), pET28b-s*xtA* ACP-AONS (*Aphanizomenon gracile*), were purchased from Twist Biosciences, with codons optimized for *E. coli* expression.

pMCSG7-*C. raciborskii sxtA* ACP-AONS genomic DNA sequence (Ala722-His1245)

GenBank accession ABI75094.1; linkers and 6x-His tags are shown in red. The serine codon in bold was mutated to alanine in the ACP knockout (ACP(KO)-AONS) construct.

atgcaccatcatcatcattcttctggtgtagatctgggtaccgagaacctgtacttccaatccaatgcgatgGC
AATTGGTTCATTGGTACCAAAAGCAACATCTGCAACTAAGGAAAACAAAACTGTAGCGGATCTCGTTAAAGAATGCA
TCTTAAAAGTAATGGGTTCCCAACGTCAGGCAGCCTACGCTCCACAACAAAAACTGCTGGATATGGGATTAGAT**TCT**
TTAGATTTATTAGAACTGCAAACGCTCCTAGAGGAACGTTTAGGGATCAATCTGTCTGGAACGTTCTTTTTACAAAA
GAACACTCCAACTGCCATCATCACTTATTTCCAAAACCAAGTGGTACAAGAGAAACAATCTGATCTAGCTCCACCTG
TTGACTCAGCCAACGAAATCAACACTCTGGAAAACGTAGTTAACCAACAAAAAATTCCTCAAGTCACAAGAGTCGTC
ACAGAACAACAAGGTCGCAAGGTGCTAATTGACGGACATTGGGTGATAGACTTTGCTTCTTGCAACTATTTAGGTCT
TGACTTGCATCCAAAAGTTAAGGAAGCAATTCCACCAGCTTTGGATAAATGGGGCACACATCCAAGCTGGACTCGGC
TTGTTGCTTCCCCAGCAATTTATGAGGAATTGGAGGAAGAATTGTCCAAACTTTTAGGCGTACCTGATGTTTTAGTA
TTTCCAGCTGTAACACTGCTTCAGATAGGAATTTTACCACTATTAACTGGGAATAATGGTGTCATCTTTGGTGACAT
AGCTGCACATCGTTGTATTTATGAAGCGTGCTGTCTGGCTCAGCACAAAGGAGCCCAGTTCATCCAATATCGACATA
ATGATTTGAACGATTTAGCCGAAAAACTAGCAAAATATCCGCCTGAACAAGTAAAGATTATTGTCATTGATGGCGTG
TATTCCATGTCGGCAGATTTTCCCGATCTGCCAGCTTACGTGCATCTGGCAAAAGAGTACAATGCCTTAATTTACAT
GGATGATGCTCATGGTTTTGGCATTTTGGGCGAAAATCCCAGCAGCGATATGCCTTACGGTTACAAAGGAAACGGGA
TGGTGAATTATTTTGACCTGCGGTTTGCAGAGGATAATATCATCTATGTAGCTGGTTTGTCCAAAGCCTATTCTTCT
TACGCAGCATTCTTAACTTGTGGCGATCGCCGGATCAAAACCAACTTCCGCAACGCTTGGACTGCCATATTTTCTGG
TCCTTCTCCTGTTGCGAGTTTGGCAAGTGCCTTAGCCGGATTACAGGTGAATCGTCAGGAGGGGGAGCAGTTAAGAA
AACAAATTTATCACCTAACTCACAAATTGGTTACACAAGCAAGAGCCATTGGATTCGAAGTGGATAACTATGGTTAC
GTTCCCATCGTAGGCGTGTTAGTGGGAGATGCTCAACACATGATTGATGTGTGTCAACTCCTTTGGGAATATGGTAT
TTTAATTACTCCTGCTATTTTTCCAATCGTACCTTTAAATAAAAGTGCTTTAAGGTTTTCGATTACAGCCGCCAATA
CCGAAGAGGAGATAGACCAAGCAATTAAATCTCTCAAAGCAGTTTGGGATTTGCTACAAAAAAGGAAAGCTTTGCCT
TGTAAGCAGGAGGAAAACATACTCAAGCAT

Translation

MHHHHHHSSGVDLGTENLYFQSNAMAIGSLVPKATSATKENKTVADLVKECILKVMGSQRQAAYAPQQKLLDMGLDS
LDLLELQTLLEERLGINLSGTFFLQKNTPTAIITYFQNQVVQEKQSDLAPPVDSANEINTLENVVNQQKIPQVTRVV
TEQQGRKVLIDGHWVIDFASCNYLGLDLHPKVKEAIPPALDKWGTHPSWTRLVASPAIYEELEEELSKLLGVPDVLV
FPAVTLLQIGILPLLTGNNGVIFGDIAAHRCIYEACCLAQHKGAQFIQYRHNDLNDLAEKLAKYPPEQVKIIVIDGV
YSMSADFPDLPAYVHLAKEYNALIYMDDAHGFGILGENPSSDMPYGYKGNGMVNYFDLRFAEDNIIYVAGLSKAYSS
YAAFLTCGDRRIKTNFRNAWTAIFSGPSPVASLASALAGLQVNRQEGEQLRKQIYHLTHKLVTQARAIGFEVDNYGY
VPIVGVLVGDAQHMIDVCQLLWEYGILITPAIFPIVPLNKSALRFSITAANTEEEIDQAIKSLKAVWDLLQKRKALP
CKQEENILKH

pMCSG7-*C. raciborskii sxtA* ACP genomic DNA sequence (Ala722-Gln815)

atgcaccatcatcatcattcttctggtgtagatctgggtaccgagaacctgtacttccaatccaatgcgatgGC
AATTGGTTCATTGGTACCAAAAGCAACATCTGCAACTAAGGAAAACAAAACTGTAGCGGATCTCGTTAAAGAATGCA
TCTTAAAAGTAATGGGTTCCCAACGTCAGGCAGCCTACGCTCCACAACAAAAACTGCTGGATATGGGATTAGATTCT
TTAGATTTATTAGAACTGCAAACGCTCCTAGAGGAACGTTTAGGGATCAATCTGTCTGGAACGTTCTTTTTACAAAA
GAACACTCCAACTGCCATCATCACTTATTTCCAAAACCAAGTGGTACAA

Translation
MHHHHHHSSGVDLGTENLYFQSNAMAIGSLVPKATSATKENKTVADLVKECILKVMGSQRQAAYAPQQKLLDMGLDS
LDLLELQTLLEERLGINLSGTFFLQKNTPTAIITYFQNQVVQ

pET28b-*A. gracile sxtA* ACP-AONS *E. coli*-optimized DNA sequence (Ala735-His1234)

GenBank accession LT549448.1; linkers and 6x-His tags are shown in red.

atgggcagcagccatcatcatcatcacagcagcggcctggtgccgcgcggcagccatatggctagcatgactgg
tggacagcaaatgggtcgcggatccgaattcGCAACATCTGCGACTAAGGAGAACAAGACCGTAGCTGCCTTAGTCA
AAGAATGTGTGCTTAAAGTGATGGGCAGCCAGCGCGAGGCCGCTTATGCCCCTGAACAAAAATTGTTAGACATGGGG
TTAGATAGTTTAGACCTGTCTGAACTTCAGACTCTTTTGGAAAAACGTTTGGGCGTCAATCTTTCTGGGACTTTCTT
TTTTCAAAACAATACCCCCACTGCTATCATCACATACTTTCAGAACCAAGTTGTCCAAGAAAAGCAGAGTGATCTGG
CGAGCCCCGTGGACTTAGCTAATGAAATCAACACATTAGAAAATGTTGTCAATCAGCAGAAGATCCCTCAGGTTACG
CGCGTAGTGACTGAACAGCAAGGTCGTAAGGTCTTGATCGATGGCCACTGGGTGATTGACTTCGCCTCATGCAACTA
TTTAGGTTTAGACTTGCACCCCAAAGTGAAGGAGGCTATCCCCCCGGCTTTGGAGAAGTGGGGGGTCCATCCCTCTT
GGACACGTTTGGTGGCCAGCCCGGTAATCTATGAGGAATTAGAGGAGGAGTTAAGCAAACTGTTAGGAGTTCCCGAC
GTATTAGTTTTCCCAAGCGTCACATTACTTCAGATCGGGGTCTTGCCACTTCTGACAGGCTCGAACGGAGTTATTTT
AGGCGACATTGCTGCGCATCGTTGCATTTACGAGGCCTGCTGCCTGGCACAGCAGAAGGGCTCACAGTTTATCCAAT
ATCGCCACAATGACTTGAATGATTTAGCGGAGAAATTGGCAAAGTATCCACTGGAACAAGTCAAAATCATCGTCATT
GACGGAGTGTACTCTATGTCCGCCGACTTCCCAGACCTTCCGGCTTACGTCCGCCTGGCGAAGGAATACAATGCTTT
TATTTATATGGATGACGCGCACGGATTTGGTATTTTGGGTGAGAACCCGTCCTCCGATATGCCTTACGGATACAAAG
GTAATGGTATGGTGAATTATTTCGATCTTCAATTTGCCGAGGACAACATTATCTACGTCGCGGGGTTGTCTAAAGCC
TACAGTTCGTACGCCGCCTTTCTTACTTGCGGTGATCGTCAGATTAAGACCAATTTCCGTAATGCTTGGACTGCCAT
TTTCAGCGGTCCTAGCCCAGTGGCAAGCTTGGCATCAGCCTTGGCCGGCTTACAGGTTAATCGTCAGGAAGGAGAAC
AACTTCGCAAGCAAATCTACCACTTGACGCGCAAATTGGTGACCCAAGCTCGCGCCATTGGCTTCGAAGTGGACAAC
TACGGTTACGTTCCCATCGTTTCCGTCTTAGTCGGAGACGCGCAACACATGATTGACGTATGTCAGCTGCTGTGGGA
GTACGGCATTCTGATCACACCCGCTATTTTCCCAATTGTTCCACTTAATAAATCCGCGCTGCGCTTTTCTATCACAG
CTGCCAATACGGAAGAAGAGATTGACCAAGCCATCAAAGCATTGGAGGGGGTATGGGTCCTGCTTCAGAAAGAGAAG
gcggccgcactcgagcaccaccaccaccaccac

Translation

MGSSHHHHHHSSGLVPRGSHMASMTGGQQMGRGSEFATSATKENKTVAALVKECVLKVMGSQREAAYAPEQKLLDMG
LDSLDLSELQTLLEKRLGVNLSGTFFFQNNTPTAIITYFQNQVVQEKQSDLASPVDLANEINTLENVVNQQKIPQVT
RVVTEQQGRKVLIDGHWVIDFASCNYLGLDLHPKVKEAIPPALEKWGVHPSWTRLVASPVIYEELEEELSKLLGVPD
VLVFPSVTLLQIGVLPLLTGSNGVILGDIAAHRCIYEACCLAQQKGSQFIQYRHNDLNDLAEKLAKYPLEQVKIIVI
DGVYSMSADFPDLPAYVRLAKEYNAFIYMDDAHGFGILGENPSSDMPYGYKGNGMVNYFDLQFAEDNIIYVAGLSKA
YSSYAAFLTCGDRQIKTNFRNAWTAIFSGPSPVASLASALAGLQVNRQEGEQLRKQIYHLTRKLVTQARAIGFEVDN
YGYVPIVSVLVGDAQHMIDVCQLLWEYGILITPAIFPIVPLNKSALRFSITAANTEEEIDQAIKALEGVWVLLQKEK
AAALEHHHHHH

pET28b-*D. circinale sxtA* ACP-AONS *E. coli*-optimized DNA sequence (Ala735-His1234)

GenBank accession DQ787201.1; linkers and 6x-His tags are shown in red.

atgggcagcagccatcatcatcatcacagcagcggcctggtgccgcgcggcagccatatggctagcatgactgg
tggacagcaaatgggtcgcggatccgaattcGCTACAAAGGAAAATAAGACAGTAGCGGCTCTGGTCAAGGAGTGTG
TGCTTAAAGTCATGGGTTCCCAGCGTGAAGCTGCGTATGCCCCCGAACAAAAGCTGCTGGATATGGGCCTTGACTCC
CTTGACTTGAGTGAACTGCAGACTCTTTTAGAAAAACGCCTTGGGGTCAACCTTTCCGGTACATTTTTTTTCCAGAA
TAATACACCAACCGCAATTATCACATACTTTCAAAACCAGGTCGTACAAGAGAAGCAGAGTGACTTAGCGAGTCCCG
TTGACTTAGCGAACGAGATCAACACGTTAGAGAATGTCGTCAACCAGCAGAAAATCCCGCAAGTCACGCGTGTGGTA
ACTGAGCAACAAGGACGTAAAGTGTTAATTGACGGACATTGGGTCATTGACTTTGCGTCTTGTAATTATTTAGGACT
GGATTTGCATCCCAAGGTAAAAGAGGCGATTCCCCCAGCTCTTGAGAAGTGGGGGGTACATCCCAGTTGGACTCGTC
TGGTTGCCTCCCCAGTTATTTATGAAGAGTTGGAGGAAGAGTTGAGTAAGCTGCTGGGCGTTCCTGATGTATTAGTC
TTTCCTTCGGTGACTCTGTTGCAGATCGGTGTCTTGCCACTTTTGACGGGTAGCAACGGGGTAATTCTGGGGGATAT
TGCGGCCCATCGTTGCATCTACGAGGCCTGTTGCCTGGCTCAGCAGAAAGGGTCGCAGTTCATCCAGTACCGTCACA
ACGATTTAAATGACTTAGCCGAGAAACTTGCGAAGTACCCTCTTGAACAAGTCAAGATTATTGTGATCGACGGAGTG
TACTCTATGTCTGCGGACTTTCCAGACCTGCCGGCCTATGTTCGTTTGGCAAAAGAGTACAACGCTTTTATCTACAT
GGACGACGCACATGGTTTCGGGATCTTGGGAGAAAACCCCTCGTCAGATATGCCATACGGTTATAAAGGAAATGGTA
TGGTCAACTATTTCGACCTTCAATTCGCGGAGGACAATATCATTTATGTCGCTGGCTTATCGAAGGCATATTCCTCG
TATGCAGCATTTTTGACCTGCGGAGACCGCCAAATCAAGACAAACTTCCGCAACGCATGGACAGCTATCTTCTCCGG
TCCTTCGCCGGTAGCGTCTCTGGCATCGGCCTTAGCCGGTTTACAAGTGAATCGCCAAGAGGGCGAACAGTTGCGCA
AACAAATCTATCATCTTACTCGCAAATTGGTCACGCAGGCGCGCCATTGGATTCGAAGTAGACAACTATGGGTAT
GTCCCTATCGTGTCGGTGTTGGTAGGAGATGCCCAGCACATGATCGATGTGTGTCGCTTGCTTTGGGAATATGGCAT
CCTGATTACGCCAGCCATCTTTCCGATTGTTCCGCTGAATAAGTCTGCCTTACGTTTTTCCATCACTGCAGCAAATA
CGGAGGAAGAAATTGACCAAGCGATCAAGGCCTTAGAAGCTGTATGGGTCCTTTTACAGAAAGAAAAGgcggccgca
ctcgagcaccaccaccaccaccac

Translation

MGSSHHHHHHSSGLVPRGSHMASMTGGQQMGRGSEFATKENKTVAALVKECVLKVMGSQREAAYAPEQKLLDMGLDS
LDLSELQTLLEKRLGVNLSGTFFFQNNTPTAIITYFQNQVVQEKQSDLASPVDLANEINTLENVVNQQKIPQVTRVV
TEQQGRKVLIDGHWVIDFASCNYLGLDLHPKVKEAIPPALEKWGVHPSWTRLVASPVIYEELEEELSKLLGVPDVLV
FPSVTLLQIGVLPLLTGSNGVILGDIAAHRCIYEACCLAQQKGSQFIQYRHNDLNDLAEKLAKYPLEQVKIIVIDGV
YSMSADFPDLPAYVRLAKEYNAFIYMDDAHGFGILGENPSSDMPYGYKGNGMVNYFDLQFAEDNIIYVAGLSKAYSS
YAAFLTCGDRQIKTNFRNAWTAIFSGPSPVASLASALAGLQVNRQEGEQLRKQIYHLTRKLVTQARAIGFEVDNYGY
VPIVSVLVGDAQHMIDVCRLLWEYGILITPAIFPIVPLNKSALRFSITAANTEEEIDQAIKALEAVWVLLQKEKAAA
LEHHHHHH

pET28b-*M. wollei sxtA* ACP-AONS *E. coli*-optimized DNA sequence (Ala732-His1243)

GenBank accession EU603711.1; linkers and 6x-His tags are shown in red.

atgggcagcagccatcatcatcatcacagcagcggcctggtgccgcgcggcagccatatggctagcatgactgg
tggacagcaaatgggtcgcggatccgaattcGCTACCAAAGAGAACGAAACTGTAGCAGCGTTAGTAAAAGAGTGTT
TCTTGAAAGTTCTGGGATCGCAGCGTCAAGCAGCTTATGCGCCAAAGCAGAAATTGTTGGACATGGGCCTGGACAGC
CTTGACTTGCTTGAGCTGCAGACACTGCTTGAAGAGCGCCTTGGAGTGAGTCTGTCACGTACGTTCTTCTTGCAGAA
TAATACCCCCACAGCCATTATCACCTACTTCCAAACATATTTCCAGAATCAGGTAGTTCAAGAAAAGAAGTCCGATC
TGGCATCCCCCGTGGACTCCGCTAACGAAATTAATACCTTAGAGAATGTAGTGAACCAGCAAAAGATCCCTCAAGTT
ACCCGCGTTGTGACCGAACAGCAAGATCGCAAGGTTTTAATCGACGGCCATTGGGTGATCGATTTCGCCTCCTGCAA
TTACTTAGGATTAGACTTGCATCCCAAAGTCAAAGAGGCCATCCCCCCTGCGTTGGAAAAGTGGGGCGTACACCCAT
CTTGGACCCGCTTGGTAGCATCTCCAGCAATCTATGAGGAGCTGGAAGAGGAGCTTGCAAAACTTTTAGGGGTACCA
GACGTATTGTTGTTCCCATCCTTGACCTTACTGCAGATGGGGGTACTGCCCCTGCTGACCGGCAATAATGGAGTGAT
CTTCGGGGACATCTCCGCCCATCGCTGTATCTATGAGGCGTGCTGCTTAGCACAACATAAAGGAGCGCAATTCATTC
AGTACCGTCATAACGATCTGAATGACTTGGCGGCGAAATTGGCAAAATACCCGCTGGAACAAGTGAAGATCATCGCG
ATTGATGGGGTGTACAGCATGAGTGCTGACTTGCCAGACTTGCCCGCGTATGTCCGCTTAGCAAAACAGTATAATGC
TTTGATCTATATCGACGATGCTCATGGGTTCGGAATCCTTGGGGAAAATCCATCGAGTGATATGCCTTATGGATACA
AAGGAAACGGTATCGTAAATTATTTCGATCTTCGCTTCGCTGAAGACAACATCATTTACGTTGCTGGTCTTTCCAAA
GCCTACAGTTCATACGCAGCATTCGTTACTTGCGGAAATTCCCAAATCAAGACGCAGTTTCGTAACGCGTGGACCGC
CATTTTTTCTGGACCGTCACCAGTTGCATCGTTGGCAAGCGCCCTTGCTGGTCTTCAGGTCAATCGCCAAGAGGGGG
AGCAATTGCGTAAGCAGGTTTATTACTTAACGCATAAGTTAGTGACCCAGGCCCGCGCGATTGGTTTTGAAGTAGAC
AACTACTCATACGTGCCGATTGTCTCAGTACTTGTAGGGGACGCCCAGCATATCGTTGACGTCTGCCAGTTGTTATG
GGAATACGGAATTTTAATTACACCCGCGATTTTCCCTATTGTTCCGCTGAATAAAACCGCATTGCGTTTCAGCATTA
CCGCTGCGAACACCGAGGAGGAAATTGACCAAGCGATCAAAGCACTGGAGGCTGTATGGGACTTGCTGCAAAAGCGC
AAAGCCCTTTTATGTAAGgcggccgcactcgagcaccaccaccaccaccac

Translation

MGSSHHHHHHSSGLVPRGSHMASMTGGQQMGRGSEFATKENETVAALVKECFLKVLGSQRQAAYAPKQKLLDMGLDS
LDLLELQTLLEERLGVSLSRTFFLQNNTPTAIITYFQTYFQNQVVQEKKSDLASPVDSANEINTLENVVNQQKIPQV
TRVVTEQQDRKVLIDGHWVIDFASCNYLGLDLHPKVKEAIPPALEKWGVHPSWTRLVASPAIYEELEEELAKLLGVP
DVLLFPSLTLLQMGVLPLLTGNNGVIFGDISAHRCIYEACCLAQHKGAQFIQYRHNDLNDLAAKLAKYPLEQVKIIA
IDGVYSMSADLPDLPAYVRLAKQYNALIYIDDAHGFGILGENPSSDMPYGYKGNGIVNYFDLRFAEDNIIYVAGLSK
AYSSYAAFVTCGNSQIKTQFRNAWTAIFSGPSPVASLASALAGLQVNRQEGEQLRKQVYYLTHKLVTQARAIGFEVD
NYSYVPIVSVLVGDAQHIVDVCQLLWEYGILITPAIFPIVPLNKTALRFSITAANTEEEIDQAIKALEAVWDLLQKR
KALLCKAAALEHHHHHH

Table 3.S1. Primers used to generate constructs in this Chapter.

Name	Sequence
Primers for SxtA(S773A) by site-directed mutagenesis of whole plasmids	
SxtA S773A forward	5′-AAACTGCTGGATATGGGATTAGATGCGTTAGATTTATTAGAACTGCAAACG-3′
SxtA S773A reverse	5′-CGTTTGCAGTTCTAATAAATCTAACGCATCTAATCCCATATCCAGCAGTTT-3′
Primers for ligation-independent cloning (LIC)	
SxtA ACP A722 LIC forward	5′-**TACTTCCAATCCAATGCGATG**GCAATTGGTTCATTGGTAC-3′
SxtA ACP Q815 LIC reverse	5′-**TTATCCACTTCCAATGTTA**TTGTACCACTTGGTTTTGGAAA-3′
SxtA AONS H1245 LIC reverse	5′-**TACTTCCAATCCAATGCAATG**CAAGAGAAACAATCTGATCT-3′
Primers for megaprimer PCR of whole plasmids (MEGAWHOP)	
MW SxtA AONS mega forward	5′-CCAAACATATTTCCAGAATCAGGTAGTTC-3′
MW SxtA AONS mega reverse	5′-GTGCTCGAGTGCGGCCGC-3′

Overhangs for ligation-independent cloning are shown in boldface, mutations underlined

***C. raciborskii* ACP knockout SxtA(S773A):** pET28b-*sxtA*(S773A) was generated by site-directed mutagenesis on pET28b-*sxtA*(wt). 50 µL PCR reaction mixtures contained 10 µL Phusion HF buffer, 2 ng/µL wt parent plasmid, 2 µM each of the S773A forward and reverse primers, 200 µM each of dNTPs, 0.04 U/µL Phusion HF and 6% (v/v) DMSO. Amplification was accomplished with the following PCR procedure: 95 °C for 2:00, (95 °C 0:30, 60 °C 1:00, 68 °C 6:00) for 18 cycles, with a final extension of 68 °C for 15:00. This was followed by a 10 µL digestion containing 1 µL NEB CutSmart buffer, 8 µL of the PCR mixture and 20 units of DpnI. The reaction mixture was incubated at 37 °C for 3 h and transformed into chemically competent DH5α cells.

***C. raciborskii* SxtA ACP:** The excised domain was amplified from pET28b-*sxtA*(wt) in 50 µL reactions containing 10 µL Phusion HF buffer, 2 ng/µL parent plasmid, 2 µM each of forward and reverse LIC primers, 200 µM each of dNTPs, 0.04 U/µL Phusion HF. Excised domains were amplified according to the following PCR procedure: 95 °C 2:00, (95 °C 0:30, 46-50 °C 0:30, 72 °C 0:45/kb) for 30 cycles, 72 °C 10:00. Inserts were purified by gel extraction and subcloned into pMCSG7 using standard LIC protocols.[88]

***C. raciborskii* ACP-AONS and ACP(KO)-AONS**: The excised didomains were amplified from pET28b-*sxtA* (*Cylindrospermopsis raciborskii* T3) and pET28b-*sxtA*(S773A) in 50 µL reactions containing 10 µL Phusion HF buffer, 2 ng/µL parent plasmid, 2 µM each of forward and reverse LIC primers, 200 µM each of dNTPs, 0.04 U/µL Phusion HF. The didomain was amplified according to the following PCR procedure: 95 °C 2:00, (95 °C 0:30, 55 °C 1:00, 72 °C 2:00) for 30 cycles, 72 °C 10:00. The insert was purified by gel extraction and subcloned into pMCSG7 using standard LIC protocols.[88]

Library generation by megaprimer PCR of whole plasmids (MEGAWHOP): The megaprimer was amplified from pET28b-*sxtA* (*Microseira wollei*) in 100 µL reactions containing 10 µL Taq buffer, 2 ng/µL parent plasmid, 1.5 µM each of forward and reverse MEGAWHOP primers, 20 µM dGTP, 20 µM dATP, 100 µM dCTP, 100 µM dTTP, 0.2 U/µL Taq polymerase and 3 mM $MgCl_2$. The AONS domain was amplified according to the following PCR procedure: 95 °C 2:00, (95 °C 0:20, 52 °C 0:30, 72 °C 1:30) for 30 cycles, 72 °C 10:00. The megaprimer was purified by gel extraction and then used to amplify whole plasmids in 50 µL reactions containing 10 µL Phusion HF buffer, 0.5 ng/µL parent plasmid, 10 ng/µL megaprimer, 200 µM each of dNTPs, 0.04 U/µL Phusion HF polymerase. The following touchdown PCR procedure was followed: 68 °C 5:00, 98 °C 2:00, (98 °C 0:30, 68-55 °C 1:00, 68 °C 5:30) for 30 cycles with the annealing temperature decreasing by 1 °C for each of the first thirteen cycles, 72 °C 15:00. This was followed by a 10 µL digestion containing 1 µL NEB CutSmart buffer, 8 µL of the PCR mixture and 20 units of DpnI for 3 h at 37 °C, and then ligation with 1,000 U T4 DNA ligase with an additional 1.2 µL of ligase buffer at 18 °C for 12 h. The reaction mixtures were then used to transform BAP1 pGro7 *E. coli*.[71]

D. Protein overexpression and purification

Overexpression of SxtA ACP(KO)-AONS: pMCSG7 plasmids containing the desired insert were transformed into pRARE-containing chemically competent BL21(DE3) *E. coli* cells. A single colony was picked to inoculate a 5 mL LB starter culture grown overnight at 37 °C, 200 rpm. The following day, 0.5 L TB media, supplemented with spectinomycin and ampicillin, was inoculated with the starter culture and incubated at 37 °C, 250 rpm until the OD_{600} reached 1.0. Cultures were equilibrated at 20 °C for 1 h. Expression was induced by addition of IPTG (final concentration 200 μM). Cultures were incubated at 20 °C, 200 rpm for 18 h.

Overexpression of *holo*-ACP and ACP-AONS for purification: pMCSG7 or pET28 plasmids containing intact ACP-AONS were transformed into BAP1 *E. coli* cells. A single colony was picked to inoculate a 10 mL LB starter culture grown overnight at 37 °C, 200 rpm. The following day, 1 L LB media was supplemented with ampicillin or kanamycin, inoculated with the starter cultures, and incubated at 37 °C, 200 rpm until the OD_{600} reached 0.6. The culture was cooled to rt and expression was induced by addition of IPTG (final concentration 100 μM). Cultures were incubated at 18 °C, 200 rpm for 18 h.

Overexpression of ACP-AONS for whole-cell reactions: pMCSG7 or pET28 plasmids containing intact ACP-AONS were transformed into BAP1 pGro7 *E. coli* cells. A single colony was picked to inoculate a 5 mL LB starter culture grown overnight at 37 °C, 200 rpm. The following day, 0.5 L TB media was supplemented with chloramphenicol and either ampicillin or kanamycin, inoculated with the starter cultures, and incubated at 37 °C, 200 rpm until the OD_{600} reached 0.6. Arabinose (final concentration 0.5 mg/mL) was added to induce chaperone expression

while the culture cooled to 18 °C. After one hour, ACP-AONS expression was induced by addition of IPTG (final concentration 200 μM). Cultures were incubated at 18 °C, 200 rpm for 18 h. Cells were separated into approx. 0.5 g chunks and stored at -80 °C.

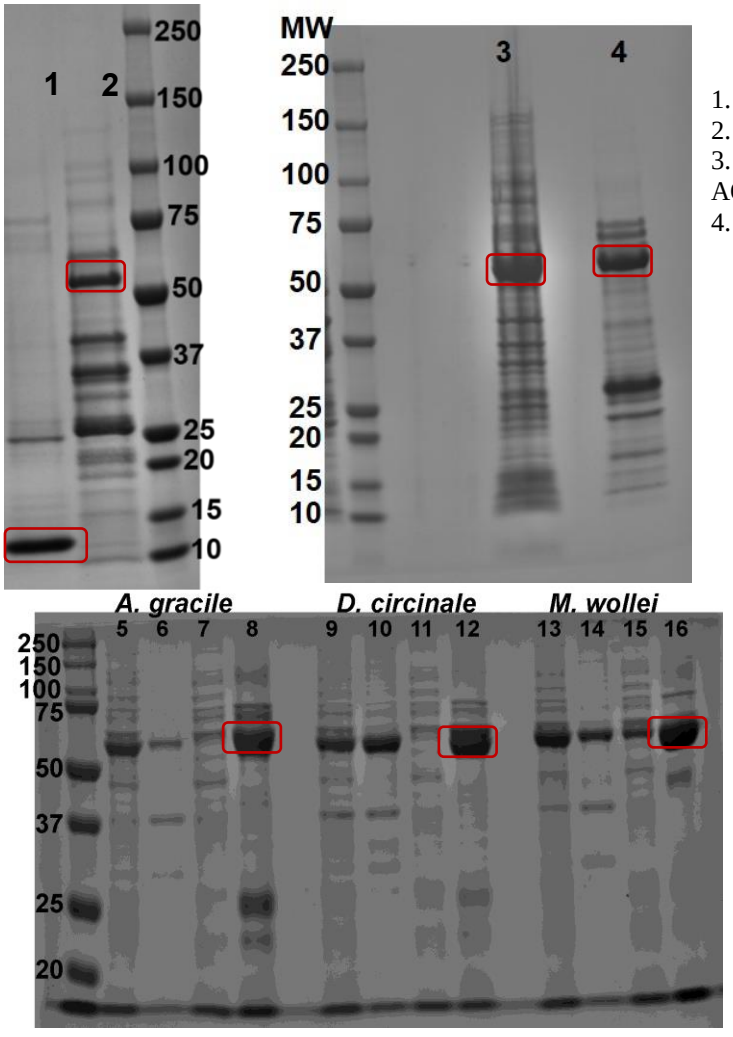
96-well plate expression of the ACP-AONS library: From an agar plate of SxtA ACP-AONS library members in BAP1 pGro7 *E. coli*, 94 colonies were picked to inoculate a 96-well plate. The 95th well contained a parent wild-type colony and the 96th well an empty pET28b vector. Each well contained 600 μL LB media supplemented with kanamycin and chloramphenicol. The plate was covered and incubated overnight at 37 °C, 250 rpm. The following day, 50 μL of each starter culture was used to inoculate a second 96-well plate containing 1 mL of TB media per well. The plate was incubated for 4 h at 37 °C, 250 rpm. Arabinose (12 μL of 25 mg/mL for a final concentration of 0.5 mg/mL) was added to each well to induce chaperone expression, while the plate cooled over 1 h to 18 °C, still 250 rpm. Protein expression was then induced with the addition of IPTG (12 μL of 10 mM IPTG for a final concentration of 200 μM). Plates grew overnight at 18 °C, 250 rpm. The following day, cells were pelleted by centrifuging the plates at 2,000 x g for 10 min. The media was decanted, and cells were resuspended in the reaction buffer (see section III).

Purification of all SxtA constructs: Cell pellets were resuspended in 4 mL of lysis buffer (1 mM PLP, 50 mM HEPES, 300 mM NaCl, 10 mM imidazole, and 10% v/v glycerol at pH 8.0) with per gram of wet cell mass. PLP was not added to *holo*-ACP cell pellets. Protease inhibitors pepstatin A (1 μg/mL), benzamidine hydrochloride (15 μg/mL), and PMSF (1 mM) were added to the resuspension mixture before incubation at 4 °C for 20 min with gentle shaking. Cells were lysed by sonication (3 s on, 6 s off, 5 min total). Insoluble material was removed by centrifugation

(30,000 x *g* for 30 min at 4 °C). The clarified lysate was incubated with gentle shaking along with 1-2 mL of nickel-NTA resin for 1 h at 4 °C and poured into a column. The resin-bound protein was washed with 25 mL wash buffer (50 mM HEPES, 300 mM NaCl, 25 mM imidazole, 10% v/v glycerol, pH 8.0). The desired protein was eluted with 6 mL elution buffer (50 mM HEPES, 300 mM NaCl, 300 mM imidazole, 10% v/v glyercol, pH 8.0). The eluted protein was concentrated using 3 – 50 kDa centrifugal cutoff filters and exchanged into storage buffer (50 mM HEPES, 200 mM NaCl, 10% v/v glycerol, pH 7.4) using a PD-10 column. The desalted protein was concentrated further using 50 kDa centrifugal cutoff filters, aliquoted, flash frozen in liquid nitrogen and stored at -80 °C.

Table 3.S2. SxtA protein constructs and molecular weights with intact initial Met.

Construct name	Plasmid	Expression line	Calculated molecular weight (Da)
CR SxtA *holo*-ACP	pMCSG7	BAP1	13,533.2
CR SxtA ACP(KO)-AONS	pMCSG7	BL21(DE3) pRARE	60,974.0
CR SxtA ACP-AONS	pMCSG7	BAP1	61,314.1
AG SxtA ACP-AONS	pET28b	BAP1	61,205.6
DC SxtA ACP-AONS	pET28b	BAP1	60,988.5
MW SxtA ACP-AONS	pET28b	BAP1	62,041.7



1. *C. raciborskii* SxtA *holo*-ACP
2. *C. raciborskii* SxtA ACP(KO)-AONS
3. Lysed BAP1 pGro7 expressing CR ACP-AONS
4. 10 µM CR ACP-AONS
5. Crude lysate of BAP1 expressing *A. gracile* ACP-AONS
6. Soluble fraction of AG ACP-AONS
7. Insoluble fraction of AG ACP-AONS
8. AG ACP-AONS eluted from Ni-NTA resin
9. Crude lysate of BAP1 expressing *D. circinale* ACP-AONS
10. Soluble fraction of DC ACP-AONS
11. Insoluble fraction of DC ACP-AONS resin
12. DC ACP-AONS eluted from Ni-NTA
13. Crude lysate of BAP1 expressing *M. wollei* ACP-AONS
14. Soluble fraction of MW ACP-AONS
15. Insoluble fraction of MW ACP-AONS
16. MW ACP-AONS eluted from Ni-NTA resin

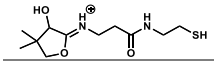
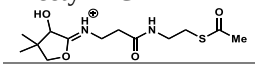
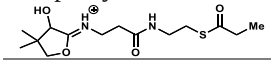
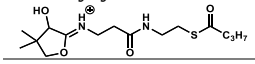
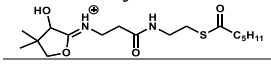
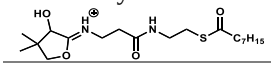
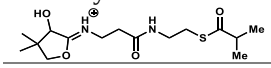
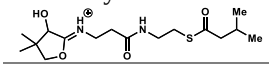
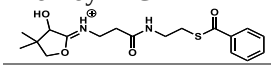
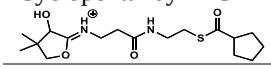
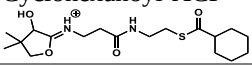
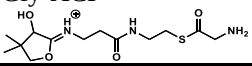
Figure 3.S1. Denatured SDS-PAGE protein gels. Proteins of interest are highlighted in red.

III. Enzymatic reactions

A. General information

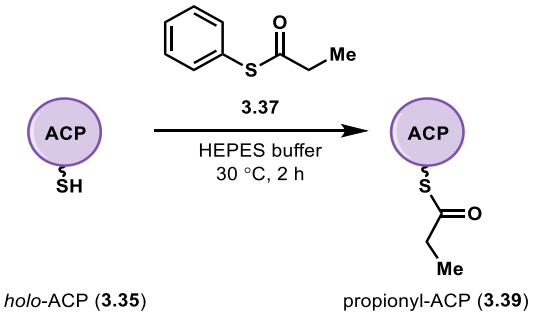
CoASH, butyryl-, isobutyryl-, isovaleryl-, octanoyl- and benzoyl-CoA were purchased from Sigma-Aldrich. Acetyl-, malonyl-, propionyl-, methylmalonyl-, and hexanoyl-CoA were obtained from CoALA Biosciences. *S*-adenosyl methionine was purchased as the tosyl salt from Carbosynth and prepared as a 50 mM aqueous solution immediately before use. Synthesized substrates and product standards were prepared as 50 mM stock solutions in DMSO. Electrospray liquid chromatography-mass spectrometry (LC-MS) analysis was performed on an Agilent G6545A quadrupole-time of flight mass spectrometer in positive mode with an Agilent 1290 UPLC system or an Agilent 6320 time of flight mass spectrometer with an Agilent 1290 Infinity II UPLC system. Solvent A = water with 0.1% formic acid. Solvent B = 95% acetonitrile, 5% water and 0.1% formic acid. MS data were analyzed using the Agilent Qualitative Mass Hunter software.

Table 3.S3. Acyl-ACP masses.

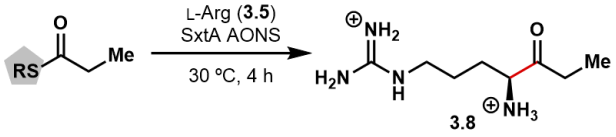
Species	Calculated intact protein mass (Da)	Observed mass (Da)[a]	Calculated Ppant ejection[89] fragment (*m/z*)	Observed Ppant ejection mass (*m/z*)
Apo-ACP	13,193.1	13,193.2	N/A	N/A
Holo-ACP 	13,533.2	13,533.2	261.13	261.13
Acetyl-ACP 	13,575.2	13,575.4	303.14	303.14
Propionyl-ACP 	13,589.2	13,589.5	317.15	317.15
n-Butyryl-ACP 	13,603.2	13,603.4	331.17	331.17
n-Hexanoyl-ACP 	13,631.3	13631.9	359.20	359.20
n-Octanoyl-ACP 	13,659.3	ND	387.23	ND
Isobutryl-ACP 	13,603.2	13,603.4	331.17	331.17
Isovaleryl-ACP 	13,617.3	13,617.8	345.18	345.18
Benzoyl-ACP 	13,637.2	13,637.9	365.15	365.15
Cyclopentanoyl-ACP 	13,629.3	13,629.4	357.18	357.18
Cyclohexanoyl-ACP 	13,643.3	ND	371.20	ND
Gly-ACP 	13,590.2	13,590.1	318.15	318.15

ND: not detected. All masses are based on *C. raciborskii* SxtA *holo*-ACP. [a]Minor species resulting from loss of the initial Met residue were occasionally observed.

B. Reactions and data



Preparation of propionyl-ACP (3.39): 1.5 mM propionyl-SPh (**3.37**), 300 µM *holo*-ACP from *C. raciborskii* SxtA in 50 mM HEPES pH 7.4, 150 mM NaCl and 10% v/v glycerol were combined (total volume 1 mL) and incubated at 30 °C for 2 h. The reaction mixture was loaded onto a PD-10 desalting column for exchange into storage buffer (50 mM HEPES, 200 mM NaCl, 10% v/v glycerol, pH 7.4). The desalted protein was concentrated further using 10 kDa centrifugal cutoff filters, aliquoted, flash frozen in liquid nitrogen and stored at -80 °C. The loading levels were analyzed by intact protein MS (about 80%) and the final protein concentration was measured by the Pierce 660 assay. LC-MS analysis: column = Phenomenex Aeris 3.6 µm WIDEPORE C4 2.1 x 50 mm; method = 5% B at 0.5 mL/min for 2 min, followed by a linear gradient to 100% B over 4 min, 100% B for 2 min, followed by a 0.1 min linear gradient to 5% B and 1.9 min equilibration at 5% B (total time 10 min). t_R = 4.1 min. The relative abundance of ejected Ppant ions for *holo*- and propionyl-ACP were used to calculate the fraction of each ACP species.[89]



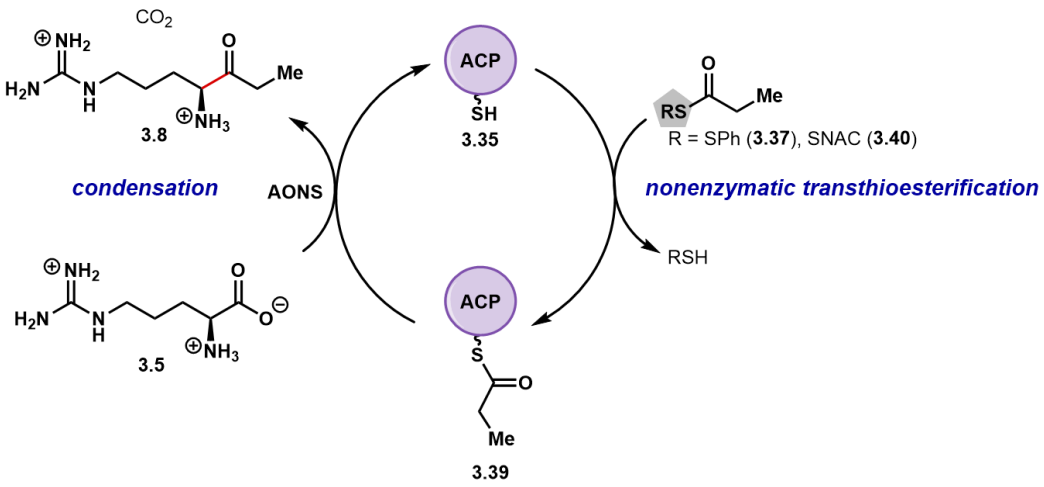
Thiol comparison (Figure 3.7):

With thiols thiophenol (SPh), 2-thiopyridine, *N*-acetylcysteamine (SNAC), pantetheine, phosphopantetheine (Ppant), 3'-dephospho-CoA, CoA and *holo*-ACP.

CoA biosynthesis. 250 μM propionyl-pantetheine (**3.38**), 2 mM Arg, 8 mM ATP, 0.2 mg/mL CoaA (to generate Ppant, see 3.32 in Figure 3.6), 0.2 mg/mL CoaD (to generate 3'-dephospho-CoA), and 0.2 mg/mL CoaE (for CoA) in 50 mM HEPES pH 7.4, 150 mM NaCl, 10 mM $MgCl_2$ and 0.5 mM $MnCl_2$ (total volume 42 μL) were combined and incubated at 30 °C for 3 h. For condensation reactions without pantetheine or the CoA biosynthetic intermediates, the volume of pantetheine and CoaADE was replaced with water. After 3 h, 21 μL aliquots were removed to check CoA biosynthesis progress and the reactions were quenched by the addition of 4 μL water and then diluting 5x with MeOH. Precipitate was pelleted at 12,000 x *g* for 20 min. The supernatant was transferred to sample vials for LC-MS analysis: column = Waters XBridge C18 3.5 μm , 2.1 x 150 mm; solvent A = water with 0.1% v/v formic acid, solvent B = methanol with 0.1% v/v formic acid; method = 10% B at 0.2 mL/min for 2 min, followed by a linear gradient to 90% B over 15 min, then a gradient to 90% over 1 min (total time 18 min). t_R (**3.38**): 11.52 min; t_R (propionyl-phosphopantetheine): 10.19 min; t_R (propionyl-3'-dephospho-CoA): 9.24 min. The CoA thioester peak was not detected, although all intermediate peaks were also no longer present.

Condensation: To the remaining 21 μL were added the remaining propionyl thioesters (SPh, SNAC, 2-thiopyridine, ACP) and 8 μM ACP(KO)-AONS for a total volume of 25 μL. The condensation reactions were incubated at 30 °C for an additional 4 h and then quenched by diluting 5x with acetonitrile. The supernatant was diluted 20x with the LC-MS analysis mixture for a final

ratio of 5% quenched reaction, 2.5% of aqueous 10 µg/mL ^{15}N Arg·HCl (Cambridge Isotope Laboratories), 46% acetonitrile, 46.5% of 1% formic acid in water (e.g., 10 µL of quenched reaction, 5 µL of 10 µg/mL ^{15}N Arg·HCl, 92 µL acetonitrile, 93 µL of 1% formic acid in water for a total volume of 200 µL). Samples were analyzed by LC-MS: column = Waters Acquity 1.7 µm UPLC BEH HILIC 2.1 x 100 mm column; method = 85% B at 0.3 mL/min for 2 min, followed by a linear gradient to 75% B at 0.4 mL/min over 3 min and a second linear gradient to 60% B over 0.5 min, 60% B for 1 min and then 6.5 min re-equilibration at 85% B at 0.3 mL/min (total time 15 min). t_R (**3.8**): 1.94 min; t_R (^{15}N Arg): 4.0 min. Yields were calculated by comparing the areas under the curve of the product extracted ion chromatograms to the ^{15}N Arg internal standard and compared to a standard curve of ketone **3.8**. Reactions were run in duplicate



Comparison of thiophenol and SNAC acyl donors and SxtA ACP in *cis* or *trans* with AONS (Figure 3.8): 2 mM propionyl thioesters, 4 mM 2-vinylpyridine, 2 mM Arg, 8 µM *C. raciborskii* ACP(KO)-AONS or ACP-AONS, 8 or 50 µM *holo*-ACP in 50 mM HEPES pH 7.4, 150 mM NaCl, 1 mM $MgCl_2$, 0.5 mM $MnCl_2$ (total volume 25 µL) were combined and incubated at 30 °C for 4 h. Reactions were quenched by diluting 5x with acetonitrile. Precipitate was pelleted at 12,000 x

g for 20 min. The supernatant was diluted 20x with the LC-MS analysis mixture for a final ratio of 5% quenched reaction, 2.5% of aqueous 10 μg/mL ^{15}N Arg·HCl (Cambridge Isotope Laboratories), 46% acetonitrile, 46.5% of 1% formic acid in water (e.g., 10 μL of quenched reaction, 5 μL of 10 μg/mL ^{15}N Arg·HCl, 92 μL acetonitrile, 93 μL of 1% formic acid in water for a total volume of 200 μL). Samples were analyzed by LC-MS: column = Waters Acquity 1.7 μm UPLC BEH HILIC 2.1 x 100 mm column; method = 85% B at 0.3 mL/min for 2 min, followed by a linear gradient to 75% B at 0.4 mL/min over 3 min and a second linear gradient to 60% B over 0.5 min, 60% B for 1 min and then 6.5 min re-equilibration at 85% B at 0.3 mL/min (total time 15 min). t_R (**3.8**): 1.94 min; t_R (^{15}N Arg): 4.0 min. Yields were calculated by comparing the areas under the curve of the product extracted ion chromatograms to the ^{15}N Arg internal standard and compared to a standard curve of ketone **3.8**. Reactions were run in duplicate.

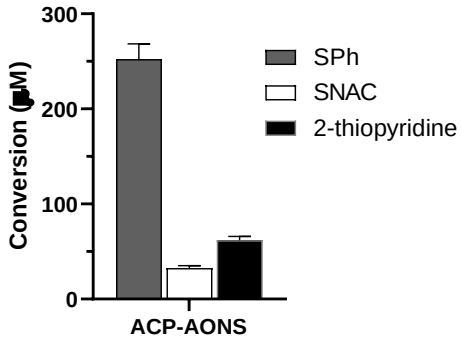
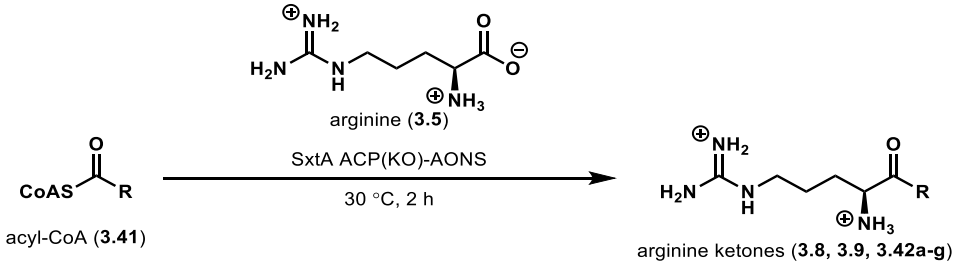
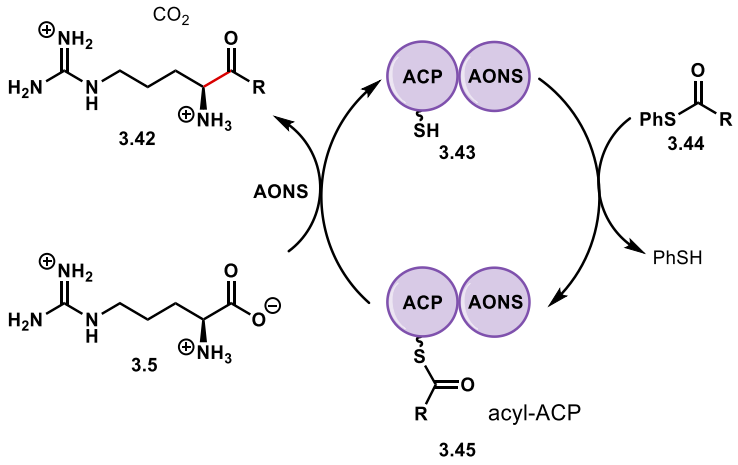


Figure 3.S2. Comparison of conversion to ketone 3.8 with different acyl donors.
The other conditions for 2-thiopyridine were not tested.

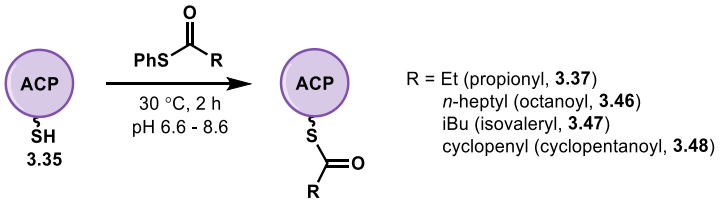


AONS-mediated condensation of arginine and acyl-CoAs (Table 3.2): 500 μM acyl-CoA, 2 mM Arg, 8 μM *C. raciborskii* ACP(KO)-AONS in 50 mM HEPES pH 7.4, 150 mM NaCl, 1 mM $MgCl_2$, 0.5 mM $MnCl_2$ (total volume 25 μL) were combined and incubated at 30 °C for 2 h. Reactions were quenched by diluting 5x with acetonitrile. Precipitate was pelleted at 12,000 x *g* for 20 min. The supernatant was diluted 20x with the LC-MS analysis mixture for a final ratio of 5% quenched reaction, 2.5% of aqueous 10 μg/mL ^{15}N Arg·HCl (Cambridge Isotope Laboratories), 46% acetonitrile, 46.5% of 1% formic acid in water (e.g., 10 μL of quenched reaction, 5 μL of 10 μg/mL ^{15}N Arg·HCl, 92 μL acetonitrile, 93 μL of 1% formic acid in water for a total volume of 200 μL). Samples were analyzed by LC-MS: column = Waters Acquity 1.7 μm UPLC BEH HILIC 2.1 x 100 mm column; method = 85% B at 0.3 mL/min for 2 min, followed by a linear gradient to 75% B at 0.4 mL/min over 3 min and a second linear gradient to 60% B over 0.5 min, 60% B for 1 min and then 6.5 min re-equilibration at 85% B at 0.3 mL/min (total time 15 min). t_R (^{15}N Arg): 4.0 min. See Supplementary Table 3.S4 for product t_R values. Yields were calculated by comparing the areas under the curve of the product extracted ion chromatograms to the ^{15}N Arg internal standard and normalized to the propionyl-CoA reaction (Table S4, entry 2). Reactions were run in triplicate.

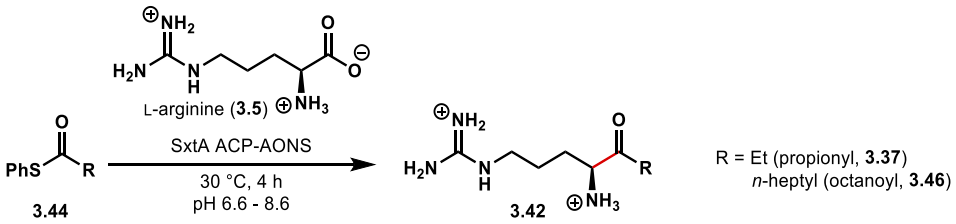


ACP-AONS-mediated condensation of arginine-based ketones, with acyl-thiophenol donors (Table 3.3): 1 mM acyl-SPh and 10 µM *C. raciborskii* ACP-AONS in 50 mM HEPES pH 7.4, 150 mM NaCl, 1 mM $MgCl_2$, 0.5 mM $MnCl_2$ (total volume 25 µL) were combined and incubated at 30 °C for 4 h. Reactions were quenched by diluting 5x with acetonitrile. Precipitate was pelleted at 12,000 x *g* for 20 min. The supernatant was diluted 20x with the LC-MS analysis mixture for a final ratio of 5% quenched reaction, 2.5% of aqueous 10 µg/mL ^{15}N Arg·HCl (Cambridge Isotope Laboratories), 46% acetonitrile, 46.5% of 1% formic acid in water (e.g., 10 µL of quenched reaction, 5 µL of 10 µg/mL ^{15}N Arg·HCl, 92 µL acetonitrile, 93 µL of 1% formic acid in water for a total volume of 200 µL). Samples were analyzed by LC-MS: column = Waters Acquity 1.7 µm UPLC BEH HILIC 2.1 x 100 mm column; method = 85% B at 0.3 mL/min for 2 min, followed by a linear gradient to 75% B at 0.4 mL/min over 3 min and a second linear gradient to 60% B over 0.5 min, 60% B for 1 min and then 6.5 min re-equilibration at 85% B at 0.3 mL/min (total time 15 min). t_R (^{15}N Arg): 4.0 min. See Supplementary Table 3.S4 for product t_R values. Yields were calculated by comparing the areas under the curve of the product extracted ion chromatograms to the ^{15}N Arg internal standard and normalized to the propionyl-CoA reaction (Table 3.2, entry 2). Reactions were run in duplicate.

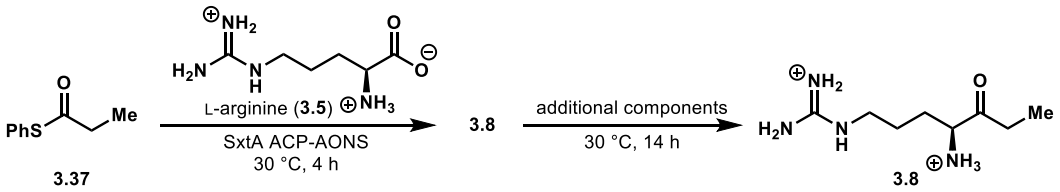
Loading: 1 mM acyl-SPh and 50 µM *C. raciborskii* SxtA *holo*-ACP in 50 mM HEPES pH 7.4, 150 mM NaCl, 1 mM $MgCl_2$, 0.5 mM $MnCl_2$ (total volume 25 µL) were combined and incubated at 30 °C for 1.5 h. The reactions were diluted to 4 µM ACP with 1% formic acid in water and transferred to sample vials for intact protein MS analysis: column = Phenomenex Aeris 3.6 µm WIDEPORE C4 2.1 x 50 mm; method = 5% B at 0.5 mL/min for 2 min, followed by a linear gradient to 100% B over 4 min, 100% B for 2 min, followed by a 0.1 min linear gradient to 5% B and 1.9 min equilibration at 5% B (total time 10 min). t_R = 4.1-4.2 min. The relative abundance of ejected Ppant ions for *holo*- and acyl-ACP were used to calculate the fraction of loaded ACP.



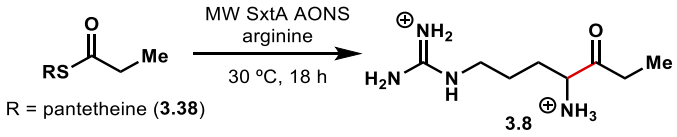
pH screen of nonenzymatic transthioesterification (Figure 3.9A): 1 mM acyl-SPh, 50 µM *C. raciborskii holo*-ACP and 4 mM Arg in 50 mM HEPES, 150 mM NaCl, 1 mM $MgCl_2$, 0.5 mM $MnCl_2$ (total volume 25 µL) were combined and incubated at 30 °C for 2 h. The pH ranged from 6.6 to 8.6, in 0.4 unit increments. The reactions were diluted to 4 µM ACP with 1% formic acid in water and transferred to sample vials for intact protein MS analysis: column = Phenomenex Aeris 3.6 µm WIDEPORE C4 2.1 x 50 mm; method = 5% B at 0.5 mL/min for 2 min, followed by a linear gradient to 100% B over 4 min, 100% B for 2 min, followed by a 0.1 min linear gradient to 5% B and 1.9 min equilibration at 5% B (total time 10 min). t_R = 4.1-4.2 min. The relative abundance of ejected Ppant ions for *holo*-and acyl-ACP were used to calculate the fraction of loaded ACP.



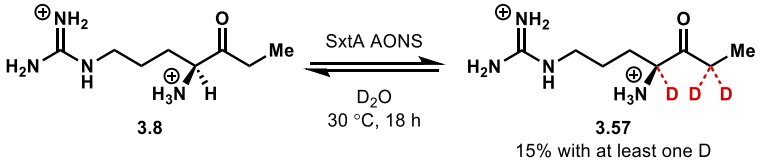
pH screen of AONS-mediated condensation (Figure 3.9B): 1 mM acyl-SPh, 4 mM Arg, and 10 µM *C. raciborskii* ACP-AONS in 50 mM HEPES, 150 mM NaCl, 1 mM $MgCl_2$, 0.5 mM $MnCl_2$ (total volume 25 µL) were combined and incubated at 30 °C for 4 h. The pH ranged from 6.6 to 8.6, in 0.2 unit increments. Reactions were quenched by diluting 5x with methanol. Precipitate was pelleted at 12,000 x *g* for 20 min. The supernatant was diluted 20x with the LC-MS analysis mixture for a final ratio of 5% quenched reaction, 2.5% of aqueous 10 µg/mL ^{15}N Arg·HCl (Cambridge Isotope Laboratories), 46% acetonitrile, 46.5% of 1% formic acid in water (e.g., 10 µL of quenched reaction, 5 µL of 10 µg/mL ^{15}N Arg·HCl, 92 µL acetonitrile, 93 µL of 1% formic acid in water for a total volume of 200 µL). Samples were analyzed by LC-MS: column = Waters Acquity 1.7 µm UPLC BEH HILIC 2.1 x 100 mm column; method = 85% B at 0.3 mL/min for 2 min, followed by a linear gradient to 75% B at 0.4 mL/min over 3 min and a second linear gradient to 60% B over 0.5 min, 60% B for 1 min and then 6.5 min re-equilibration at 85% B at 0.3 mL/min (total time 15 min). t_R (^{15}N Arg): 4.0 min. See Table 3.S4 for product t_R values. Yields were calculated by comparing the areas under the curve of the product extracted ion chromatograms to the ^{15}N Arg internal standard. Reactions were run in duplicate.



AONS-mediated condensation to identify the limiting reagent (Figure 3.10): 1 mM propionyl-SPh (**3.37**), 4 mM Arg, 10 µM *C. raciborskii* ACP-AONS in 50 mM HEPES pH 7.0, 150 mM NaCl, 1 mM $MgCl_2$, 0.5 mM $MnCl_2$ (total volume 25 µL) were combined and incubated at 30 °C for 4 h. Additional components: 1 mM propionyl-SPh, 50 µM *holo*-ACP, 10 µM ACP-AONS, 10 µM ACP(KO)-AONS or buffer were added for a final volume of 35 µL. The reactions incubated at 30 °C for 14 more hours. Reactions were quenched by diluting 5x with methanol. Precipitate was pelleted at 12,000 x *g* for 20 min. The supernatant was diluted 20x with the LC-MS analysis mixture for a final ratio of 5% quenched reaction, 2.5% of aqueous 10 µg/mL ^{15}N Arg·HCl (Cambridge Isotope Laboratories), 46% acetonitrile, 46.5% of 1% formic acid in water (e.g., 10 µL of quenched reaction, 5 µL of 10 µg/mL ^{15}N Arg·HCl, 92 µL acetonitrile, 93 µL of 1% formic acid in water for a total volume of 200 µL). Samples were analyzed by LC-MS: column = Waters Acquity 1.7 µm UPLC BEH HILIC 2.1 x 100 mm column; method = 85% B at 0.3 mL/min for 2 min, followed by a linear gradient to 75% B at 0.4 mL/min over 3 min and a second linear gradient to 60% B over 0.5 min, 60% B for 1 min and then 6.5 min re-equilibration at 85% B at 0.3 mL/min (total time 15 min). t_R (**3.8**): 1.94 min; t_R (^{15}N Arg): 4.0 min. Yields were calculated by comparing the areas under the curve of the product extracted ion chromatograms to the ^{15}N Arg internal standard and compared to a standard curve of ketone **3.8**. Reactions were run in duplicate.

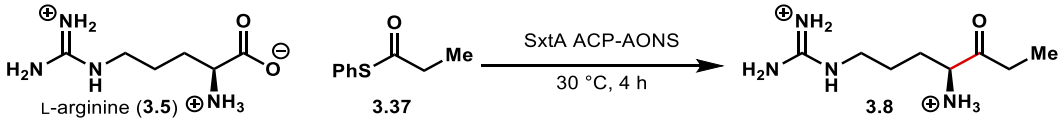


AONS-mediated condensation with both arginine enantiomers (Figure 3.11): 1 mM propionyl-pantetheine (**3.38**), 4 mM D- or L-arginine, 10 µM *M. wollei* ACP-AONS in 50 mM HEPES pH 7.0 and 150 mM NaCl (total volume) were combined and incubated at 30 °C for 18 h. At various timepoints, 10 µL aliquots were removed and quenched by diluting 5x with methanol. Precipitate was pelleted at 12,000 x *g* for 20 min. The supernatant was diluted 20x with the LC-MS analysis mixture for a final ratio of 5% quenched reaction, 2.5% of aqueous 10 µg/mL ^{15}N Arg·HCl (Cambridge Isotope Laboratories), 46% acetonitrile, 46.5% of 1% formic acid in water (e.g., 10 µL of quenched reaction, 5 µL of 10 µg/mL ^{15}N Arg·HCl, 92 µL acetonitrile, 93 µL of 1% formic acid in water for a total volume of 200 µL). Samples were analyzed by LC-MS: column = Waters Acquity 1.7 µm UPLC BEH HILIC 2.1 x 100 mm column; method = 85% B at 0.3 mL/min for 2 min, followed by a linear gradient to 75% B at 0.4 mL/min over 3 min and a second linear gradient to 60% B over 0.5 min, 60% B for 1 min and then 6.5 min re-equilibration at 85% B at 0.3 mL/min (total time 15 min). t_R (**3.8**): 1.94 min; t_R (^{15}N Arg): 4.0 min. Yields were calculated by comparing the areas under the curve of the product extracted ion chromatograms to the ^{15}N Arg internal standard and compared to a standard curve of ketone **3.8**. Reactions were run in duplicate. The inconsistent values at early timepoints were likely due to evaporation of the quenched mixtures before LC-MS analysis.



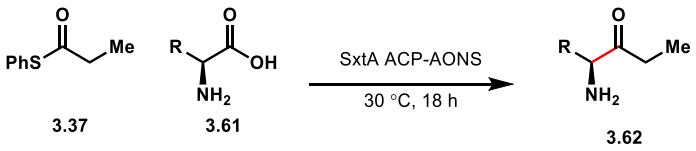
Deuterium incorporation into SxtA product ketone 3.8 (Figure 3.14A): 2 mM ketone **3.8**, 50 mM HEPES pH 7.0, 150 mM NaCl in D_2O were combined for a total volume of 25 µL and incubated at 30 °C for 18 h. The reaction was quenched by diluting 5x with methanol. Precipitate was pelleted at 12,000 x *g* for 20 min. The supernatant was diluted 20x with the LC-MS analysis mixture for a final ratio of 5% quenched reaction, 46% acetonitrile, 49% of 1% formic acid in water (e.g., 10 µL of quenched reaction, 92 µL acetonitrile, 98 µL of 1% formic acid in water for a total volume of 200 µL). Samples were analyzed by LC-MS: column = Waters Acquity 1.7 µm UPLC BEH HILIC 2.1 x 100 mm column; method = 85% B at 0.3 mL/min for 2 min, followed by a linear gradient to 75% B at 0.4 mL/min over 3 min and a second linear gradient to 60% B over 0.5 min, 60% B for 1 min and then 6.5 min re-equilibration at 85% B at 0.3 mL/min (total time 15 min). t_R (**3.8**): 1.94 min. The reactions were run in duplicate.

Chromatograms were extracted for the monoisotopic mass, *m*+1, *m*+2, m+3 (1-3 heavy atoms) for *z* = 1, and integrated. The natural abundances of heavy atom isotopologs (^{13}C, ^{15}N as calculated by enviPat https://www.envipat.eawag.ch/index.php) were subtracted from the *m*+1 and *m*+2 integration values before calculating the fraction of deuterated ketone.



Comparison of condensation mediated by different SxtA ACP-AONS homologs (Figure 3.15): 1 mM propionyl-SPh (**3.37**), 4 mM Arg, 10 µM ACP-AONS homologs (four total) in 50 mM HEPES pH 7.0, 150 mM NaCl, 1 mM $MgCl_2$, 0.5 mM $MnCl_2$ (total volume 25 µL) were

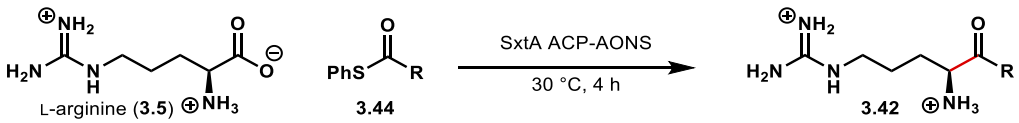
combined and incubated at 30 °C for 4 h. Reactions were quenched by diluting 5x with methanol. Precipitate was pelleted at 12,000 x *g* for 20 min. The supernatant was diluted 20x with the LC-MS analysis mixture for a final ratio of 5% quenched reaction, 2.5% of aqueous 10 μg/mL ^{15}N Arg·HCl (Cambridge Isotope Laboratories), 46% acetonitrile, 46.5% of 1% formic acid in water (e.g., 10 μL of quenched reaction, 5 μL of 10 μg/mL ^{15}N Arg·HCl, 92 μL acetonitrile, 93 μL of 1% formic acid in water for a total volume of 200 μL). Samples were analyzed by LC-MS: column = Waters Acquity 1.7 μm UPLC BEH HILIC 2.1 x 100 mm column; method = 85% B at 0.3 mL/min for 2 min, followed by a linear gradient to 75% B at 0.4 mL/min over 3 min and a second linear gradient to 60% B over 0.5 min, 60% B for 1 min and then 6.5 min re-equilibration at 85% B at 0.3 mL/min (total time 15 min). t_R (**3.8**): 1.94 min; t_R (^{15}N Arg): 4.0 min. Yields were calculated by comparing the areas under the curve of the product extracted ion chromatograms to the ^{15}N Arg internal standard and compared to a standard curve of ketone **3.8**. Reactions were run in duplicate.



3.37 3.61 3.62

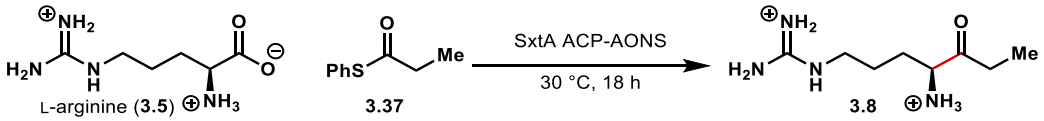
Amino acid substrate screen with ACP-AONS homologs (Figure 3.16): 1 mM propionyl-SPh (**3.37**), 4 mM L-amino acids (twenty-four total), 10 μM ACP-AONS homologs (four total) in 50 mM HEPES pH 7.0, 150 mM NaCl, 1 mM $MgCl_2$, 0.5 mM $MnCl_2$ (total volume 25 μL) were combined and incubated at 30 °C for 18 h. Reactions were quenched by diluting 5x with methanol. Precipitate was pelleted at 12,000 x *g* for 20 min. The supernatant was diluted 20x with the LC-MS analysis mixture for a final ratio of 5% quenched reaction, 2.5% of aqueous 10 μg/mL ^{15}N Arg·HCl (Cambridge Isotope Laboratories), 46% acetonitrile, 46.5% of 1% formic acid in water

(e.g., 10 µL of quenched reaction, 5 µL of 10 µg/mL ^{15}N Arg·HCl, 92 µL acetonitrile, 93 µL of 1% formic acid in water for a total volume of 200 µL). Samples were analyzed by LC-MS: column = Waters Acquity 1.7 µm UPLC BEH HILIC 2.1 x 100 mm column; method = 85% B at 0.3 mL/min for 2 min, followed by a linear gradient to 75% B at 0.4 mL/min over 3 min and a second linear gradient to 60% B over 0.5 min, 60% B for 1 min and then 6.5 min re-equilibration at 85% B at 0.3 mL/min (total time 15 min). t_R (**3.8**): 1.94 min; t_R (^{15}N Arg): 4.0 min. Yields were calculated by comparing the areas under the curve of the product extracted ion chromatograms to the ^{15}N Arg internal standard.



Acyl group substrate screen with ACP-AONS homologs (Figure 3.17): 1 mM acyl-SPh and 10 µM ACP-AONS homologs (four total) in 50 mM HEPES pH 7.0, 150 mM NaCl, 1 mM $MgCl_2$, 0.5 mM $MnCl_2$ (total volume 25 µL) were combined and incubated at 30 °C for 4 h. Reactions were quenched by diluting 5x with methanol. Precipitate was pelleted at 12,000 x *g* for 20 min. The supernatant was diluted 20x with the LC-MS analysis mixture for a final ratio of 5% quenched reaction, 2.5% of aqueous 10 µg/mL ^{15}N Arg·HCl (Cambridge Isotope Laboratories), 46% acetonitrile, 46.5% of 1% formic acid in water (e.g., 10 µL of quenched reaction, 5 µL of 10 µg/mL ^{15}N Arg·HCl, 92 µL acetonitrile, 93 µL of 1% formic acid in water for a total volume of 200 µL). Samples were analyzed by LC-MS: column = Waters Acquity 1.7 µm UPLC BEH HILIC 2.1 x 100 mm column; method = 85% B at 0.3 mL/min for 2 min, followed by a linear gradient to 75% B at 0.4 mL/min over 3 min and a second linear gradient to 60% B over 0.5 min, 60% B for 1 min and then 6.5 min re-equilibration at 85% B at 0.3 mL/min (total time 15 min). t_R (^{15}N Arg): 4.0

min. See Supplementary Table 3.S5 for product t_R values. Yields were calculated by comparing the areas under the curve of the product extracted ion chromatograms to the ^{15}N Arg internal standard. Reactions were run in duplicate. The glycyl substrate **3.S20** was not tested.



Crude format comparisons (*C. raciborskii*) and homolog whole-cell reaction comparison (Figure 3.19):

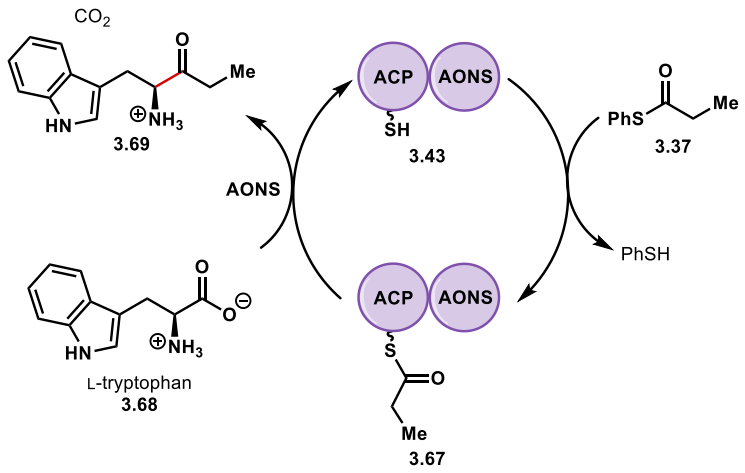
Crude formats, with BAP1 (Figure 3.18A): Cell pellets of BAP1 *E. coli* expressing *C. raciborskii* SxtA ACP-AONS (600 mg) were resuspended in 3 mL of storage buffer (50 mM HEPES, 200 mM NaCl, 10% v/v glyercol, pH 7.4) to make a 200 mg/mL suspension. The suspension was separated into a 1 mL and 2 mL portion, the latter of which was lysed by sonication (10 s on, 20 s off, 30 s total). The lysate was separated further into 2 x 1 mL portions; one portion was clarified by pelleting the debris 17,000 x *g* for 10 min at 4 °C. The biocatalytic reactions contained 1 mM propionyl-SPh (**3.37**), 4 mM Arg, 45 mg/mL of resuspended whole cells or crude lysate, or 90 mg/mL of clarified lysate in 50 mM HEPES pH 7.4, 150 mM NaCl, 1 mM $MgCl_2$, 0.5 mM $MnCl_2$, and 0.1 mM PLP (total volume 100 µL). The reactions were incubated at 30 °C for 18 h and quenched by diluting 5x with acetonitrile.

Crude formats, with BAP1 pGro7 (Figure 3.18A): Cell pellets of BAP1 pGro7 *E. coli* expressing *C. raciborskii* SxtA ACP-AONS (400 mg) were resuspended in 2 mL of storage buffer to make a 200 mg/mL suspension. The suspension was separated into 2 x 1 mL, one of which was lysed by sonication (10 s on, 20 s off, 30 s total). The biocatalytic reactions contained 1 mM propionyl-SPh (**3.37**), 4 mM Arg, 10% v/v of organic solvents (MeOH, DMSO or THF), 45 mg/mL of

resuspended whole cells or crude lysate in 50 mM HEPES pH 7.0, 150 mM NaCl, 1 mM $MgCl_2$, 0.5 mM $MnCl_2$, and 0.1 mM PLP (total volume 400 µL). The reactions were stirred at 30 °C for 18 h and quenched by diluting 5x with methanol. A crude lysate and in vitro reaction containing 10 µM of purified *C. raciborskii* ACP-AONS had comparable amounts of biocatalyst (Figure 3.18C).

Whole-cell reactions with ACP-AONS homologs (Figure 3.18B): Cell pellets of BAP1 and BAP1 pGro7 *E. coli* expressing SxtA ACP-AONS homologs (200 mg each) were resuspended in 1 mL of storage buffer. The biocatalytic reactions contained 1 mM propionyl-SPh (**3.37**), 4 mM Arg, 10% v/v of organic solvents (MeOH, DMSO or THF), 45 mg/mL of resuspended whole cells or crude lysate in 50 mM HEPES pH 7.0, 150 mM NaCl, 1 mM $MgCl_2$, 0.5 mM $MnCl_2$, and 0.1 mM PLP (total volume 400 µL). The reactions were stirred at 30 °C for 18 h and quenched by diluting 5x with methanol.

Analysis: Precipitate from quenched reactions was pelleted at 12,000 x *g* for 20 min. The supernatant was diluted 20x with the LC-MS analysis mixture for a final ratio of 5% quenched reaction, 2.5% of aqueous 10 µg/mL ^{15}N Arg·HCl (Cambridge Isotope Laboratories), 46% acetonitrile, 46.5% of 1% formic acid in water (e.g., 10 µL of quenched reaction, 5 µL of 10 µg/mL ^{15}N Arg·HCl, 92 µL acetonitrile, 93 µL of 1% formic acid in water for a total volume of 200 µL). Samples were analyzed by LC-MS: column = Waters Acquity 1.7 µm UPLC BEH HILIC 2.1 x 100 mm column; method = 85% B at 0.3 mL/min for 2 min, followed by a linear gradient to 75% B at 0.4 mL/min over 3 min and a second linear gradient to 60% B over 0.5 min, 60% B for 1 min and then 6.5 min re-equilibration at 85% B at 0.3 mL/min (total time 15 min). t_R (**3.8**): 1.94 min; t_R (^{15}N Arg): 4.0 min. Yields were calculated by comparing the areas under the curve of the product extracted ion chromatograms to the ^{15}N Arg internal standard.



Whole-cell library screening of *M. wollei* ACP-AONS variants (Figure 3.20): Cell pellets in 96-well plates were resuspended in the reaction solution containing 1 mM thioester **3.37**, 8 mM tryptophan, 50 mM HEPES pH 7.0 and 150 mM NaCl (final volume 250 µL per well). Plates were covered and incubated at 30 °C, 250 rpm for 18 h. Reactions were quenched by diluting 2x with MeOH with 0.5% v/v formic acid. Precipitate from quenched reactions was pelleted at 2,000 x *g* for 10 min. From each well 200 µL was sterile filtered through a 0.2 µm 96-well plate filter for droplet microfluidic MS analysis by Prof. Robert Kennedy's group.

Droplet Generation from Multi-well Plate (MWP)

Oil-segmented sample droplets were generated by sampling from a MWP. Quenched reaction samples were transferred from their original plate to a droplet making plate. The edges of the droplet making plate were built-up to 3 mm height using epoxy to hold oil over the wells. Samples were pulled into a 150 µm i.d. x 360 µm o.d Teflon tube (IDEX Health and Science, Oak Harbor, WA) connected to a Hamilton 25 µL gastight syringe (Reno, NV) mounted on a PHD 2000 Programmable syringe pump (Harvard Apparatus, Holliston, TX). Both the syringe and Teflon tube were prefilled with perfluorodecalin (PFD, 95% purity, Acros Organics, NJ) prior to droplet generation. A computer controlled XYZ-positioner (built in-house from XSlide assemblies, Velmex Inc., Bloomfield, NY) was used to move the inlet of the Teflon tube from sample to sample

while the syringe was withdrawing at 1.2 µL/min. Quenched reaction droplets with 45 nL volume were segmented by 50 nL oil to make a droplet train. Each train began with calibration droplets containing simulated 1 % conversion samples.

MS Analysis

Teflon tubing containing droplets was threaded through a CE-MS ESI Source equipped with coaxial sheath-flow (Agilent Technologies, Santa Clara, CA). Sheath and droplet flows were drive by Fusion 400 syringe pumps (Chemyx, Stafford, TX) at 25 µL/min and 5 µL/min respectively. Sheath buffer contained 50/50 methanol/water with 0.5% formic acid. ESI-MS analysis was performed on an Agilent 6410 triple quadrupole mass spectrometer (Agilent Technologies, Santa Clara, CA) operated in multiple reaction monitoring (MRM) mode using the parameters in Table**. ESI potential was set to 2500 V, nebulizer gas to 12 psi, and drying gas from MS source was 10 L/min at 275 °C.

Table 3.S4. Analysis of whole-cell reactions by droplet MS

Analyte	m/z transition	Dwell time (ms)	Fragmentor (V)	Collision energy (v)
Tryptophan (**3.68**)	205 → 188	150	135	15
Ketone pdt (**3.69**)	217 → 130	150	120	15

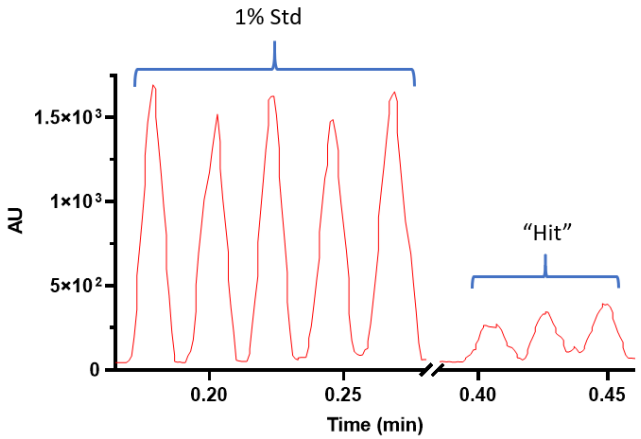
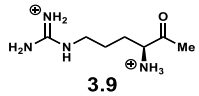
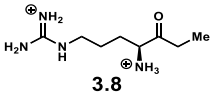
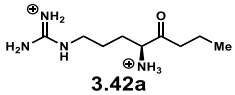
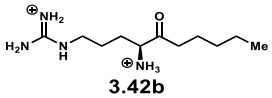
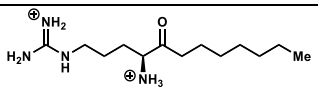
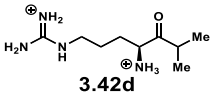
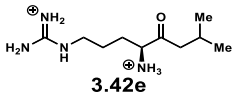
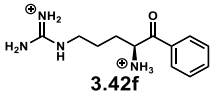
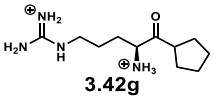
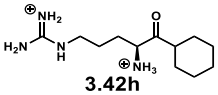


Figure 3.S3. Droplet MS of calibration droplets and potential hits of active tryptophan variants.
Data courtesy of LeeAnne Wang. Left: the 1% standard is equivalent to 1% conversion (10 µM) to ketone 3.69. Right: re-injections of the three potential hits.

Table 3.S5. Arginine-based ketone products from reactions catalyzed by the AONS domain.

Entry	R =	Thioester substrate	Product	t_R of product (min)
1	Me	acetyl-CoA or -SPh	 **3.9**	2.35
2	Et	propionyl-CoA or -SPh	 **3.8**	1.94
3	Pr	butyryl-CoA or -SPh	 **3.42a**	1.62
4	*n*-pentyl	hexanoyl-CoA or -SPh	 **3.42b**	1.24
5	*n*-heptyl	octanoyl-CoA or -SPh	 **3.42c**	1.13
6	iPr	isobutyryl-CoA or -SPh	 **3.42d**	1.67
7	iBu	isovaleryl-CoA or -SPh	 **3.42e**	1.42
8	Ph	benzoyl-CoA or -SPh	 **3.42f**	1.67
9	cyclopentyl	**3.S17**	 **3.42g**	1.58
10	cyclohexyl[a]	**3.S18**	 **3.42h**	1.41

[a]The authentic standard for this product has not been synthesized.

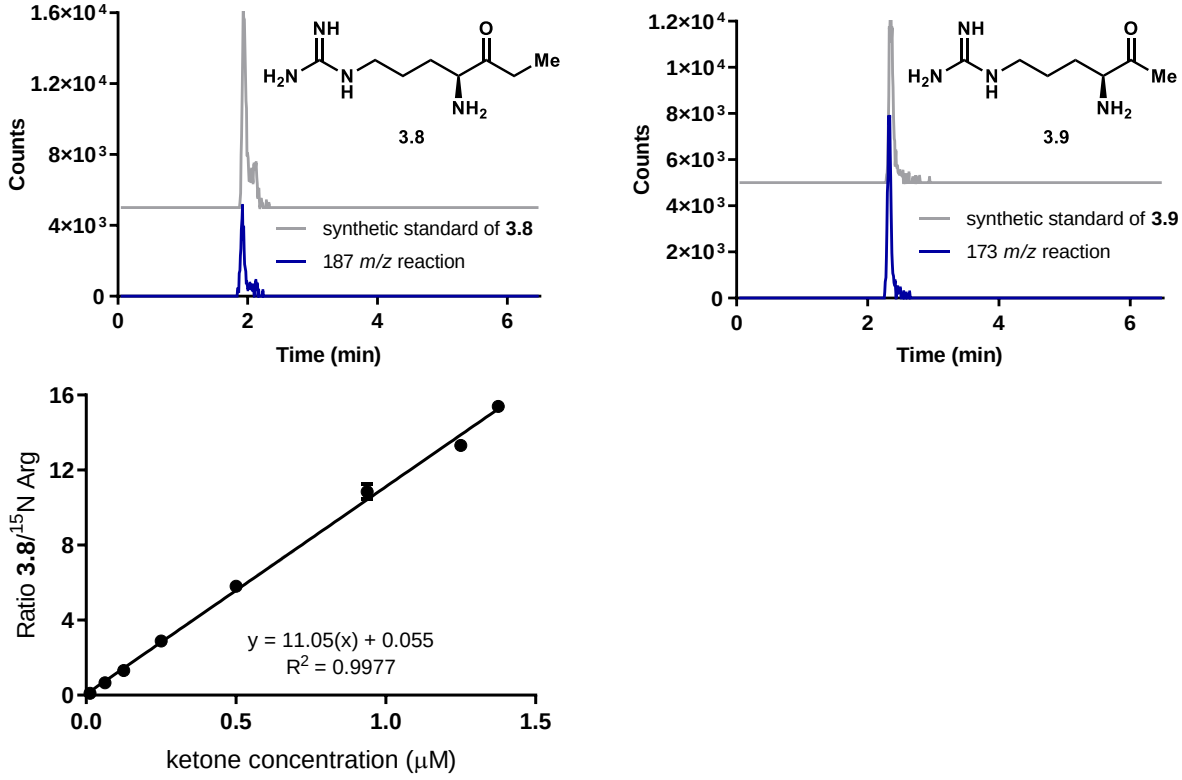


Figure 3.S4. Extracted ion chromatograms of expected arginine ketone product masses (blue) and product standards (gray) for SxtA's native product (3.8) and shunt product (3.9). A representative standard curve of ketone **3.8** used to calculate reaction conversions for AONS-mediated reactions is on the bottom left. Standard curves were not constructed for the other products, so the relative yields are compared through the ratios of area under the curves of expected product divided by the area of the internal standard, ^{15}N Arg.

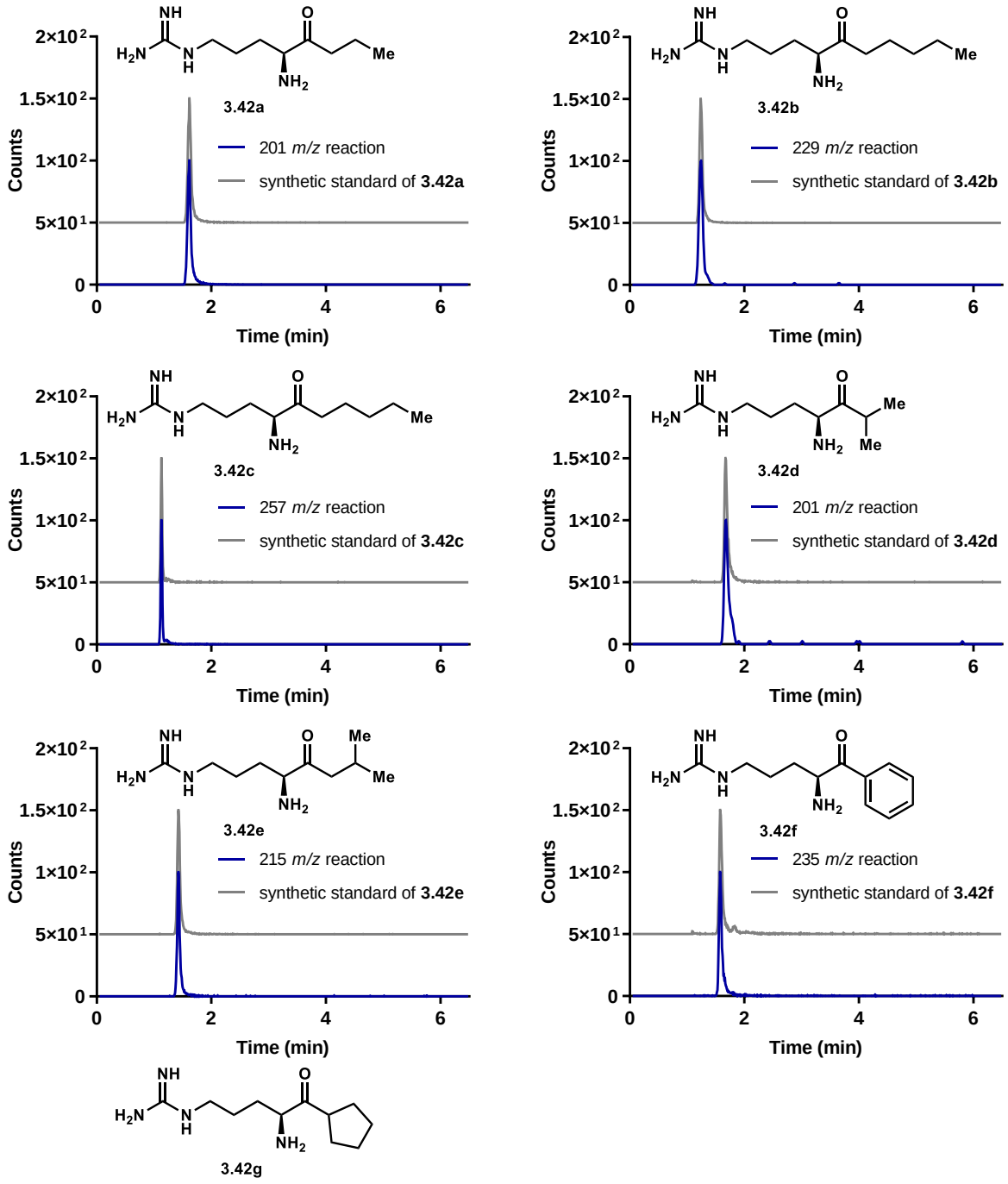
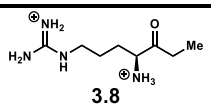
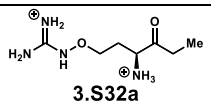
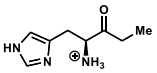
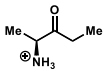
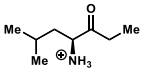
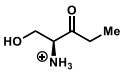
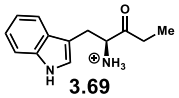


Figure 3.S5. Extracted ion chromatograms of expected arginine ketone product masses (blue) and product standards (gray).

Table 3.S6. Ethyl ketone products from reactions catalyzed by the AONS domain.

Entry	Amino acid	Product	t_R of product (min)
1	Arginine	 **3.8**	1.94
2	Canavanine	 **3.S32a**	2.14
3	Histidine[a]	 **3.S32b**	2.63
4	Glycine[a]	 **3.S32c**	1.41
5	Alanine	 **3.S24**	1.32
6	Leucine	 **3.S27**	1.14
7	Serine[a]	 **3.S32d**	1.56
8	Tryptophan	 **3.69**	1.11

[a]These products have not been verified with synthetic standards.

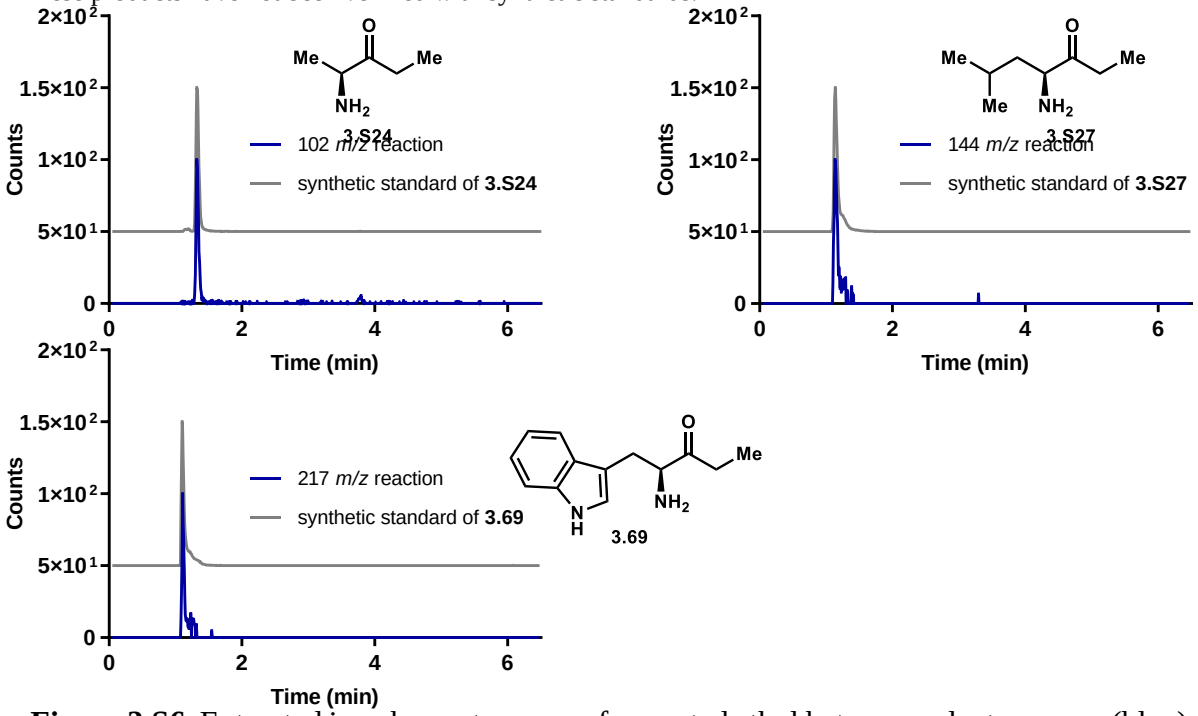


Figure 3.S6. Extracted ion chromatograms of expected ethyl ketone products masses (blue) and product standards (gray).

3.6 References

(1) Süssmuth, R. D.; Mainz, A. Nonribosomal Peptide Synthesis — Principles and Prospects. *Angew. Chem. Int. Ed.* **2017**, *56*, 3770–3821.

(2) Luo; Dong. Recent Advances in the Discovery and Biosynthetic Study of Eukaryotic RiPP Natural Products. *Molecules* **2019**, *24*, 1541.

(3) Evans, D. A.; Bartoli, J.; Shih, T. L. Enantioselective Aldol Condensations. 2. Erythro-Selective Chiral Aldol Condensations via Boron Enolates. *J. Am. Chem. Soc.* **1981**, *103*, 2127–2129.

(4) Bueno, M. P.; Cativiela, C. A.; Mayoral, J. A.; Avenoza, A. Models for the Use of .Alpha.-Amino Acids as Chiral Auxiliaries in Asymmetric Diels-Alder Reactions. *J. Org. Chem.* **1991**, *56*, 6551–6555.

(5) Vachan, B. S.; Karuppasamy, M.; Vinoth, P.; Vivek Kumar, S.; Perumal, S.; Sridharan, V.; Menéndez, J. C. Proline and Its Derivatives as Organocatalysts for Multi- Component Reactions in Aqueous Media: Synergic Pathways to the Green Synthesis of Heterocycles. *Adv. Synth. Catal.* **2020**, *362*, 87–110.

(6) de la Torre, B. G.; Albericio, F. The Pharmaceutical Industry in 2017. An Analysis of FDA Drug Approvals from the Perspective of Molecules. *Molecules* **2018**, *23*, 533.

(7) Paek, S.; Jeong, M.; Jo, J.; Heo, Y. M.; Han, Y. T. Recent Advances in Substrate-Controlled Asymmetric Induction Derived from Chiral Pool α-Amino Acids. *Molecules* **2016**, *21*, 951.

(8) Singh, P.; Samanta, K.; Das, S. K.; Panda, G. Amino Acid Chirons: A Tool for Asymmetric Synthesis of Heterocycles. *Org. Biomol. Chem.* **2014**, *12*, 6297–6339.

(9) Prabakaran, S.; Lippens, G.; Steen, H.; Gunawardena, J. Post-Translational Modification: Nature's Escape from Genetic Imprisonment and the Basis for Dynamic Information Encoding. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **2012**, *4*, 565–583.

(10) Duan, G.; Walther, D. The Roles of Post-Translational Modifications in the Context of Protein Interaction Networks. *PLoS Comput. Biol.* **2015**, *11*, 1–23.

(11) Nájera, C.; Sansano, J. M. Catalytic Asymmetric Synthesis of α-Amino Acids. *Chem. Rev.* **2007**, *107*, 4584–4671.

(12) Isidro-Llobet, A.; Álvarez, M.; Albericio, F. Amino Acid-Protecting Groups. *Chem. Rev.* **2009**, *109*, 2455–2504.

(13) Eliot, A. C.; Kirsch, J. F. Pyridoxal Phosphate Enzymes: Mechanistic, Structural, and Evolutionary Considerations. *Annu. Rev. Biochem.* **2004**, *73*, 383–415.

(14) Chun, S. W.; Hinze, M. E.; Skiba, M. A.; Narayan, A. R. H. Chemistry of a Unique Polyketide-like Synthase. *J. Am. Chem. Soc.* **2018**, *140*, 2430–2433.

(15) Heffe, W. Die STEVENS-Umlagerung von Allyl-Phenacyl-Ammoniumsalzen. *Helv. Chim. Acta* **1964**, *47*, 1289–1292.

(16) Meltzer, P. C.; Butler, D.; Deschamps, J. R.; Madras, B. K. 1-(4-Methylphenyl)-2-Pyrrolidin-1-Yl-Pentan-1-One (Pyrovalerone) Analogues: A Promising Class of Monoamine Uptake Inhibitors. *J. Med. Chem.* **2006**, *49*, 1420–1432.

(17) Meyer, R. F.; Nicolaides, E. D.; Tinney, F. G.; Lunney, E. A.; Holmes, A.; Hoefle, M. L.; Smith, R. D.; Essenburg, A. D.; Kaplan, H. R.; Almquist, R. G. Novel Synthesis of (S)-1-[5-(Benzoylamino)-1,4-Dioxo-6-Phenylhexyl]-L-Proline and Analogs: Potent Angiotensin Converting Enzyme Inhibitors. *J. Med. Chem.* **1981**, *24*, 964–969.

(18) Concellon, J.; Rodriguez-Solla, H. Synthesis and Synthetic Applications of Alpha-Amino Ketones Derived from Natural Alpha-Amino Acids. *Curr. Org. Chem.* **2008**, *12*, 524–543.

(19) Sheppard, G. S.; Florjancic, A. S. A Practical Synthesis of A-Amino Ketones via Aryllithium Addition to N-Boc-a-Amino Acids. *Synthesis (Stuttg).* **2003**, *11*, 1653–1656.
(20) Liebeskind, L. S.; Yang, H.; Li, H. A Copper-Catalyzed, PH Neutral Construction of High-Enantiopurity Peptidyl Ketones from Peptidic S -Acylthiosalicylamides in Air at Room Temperature. *Angew. Chem.* **2009**, *121*, 1445–1449.
(21) Sengupta, S.; Mondal, S.; Das, D. Amino Acid Derived Morpholine Amides for Nucleophilic α-Amino Acylation Reactions: A New Synthetic Route to Enantiopure α-Amino Ketones. *Tetrahedron Lett.* **1999**, *40*, 4107–4110.
(22) Genna, D. T.; Posner, G. H. Cyanocuprates Convert Carboxylic Acids Directly into Ketones. *Org. Lett.* **2011**, *13*, 5358–5361.
(23) Fisher, L. E.; Muchowski, J. M. Synthesis of α-Aminoaldehydes and α-Aminoketones. A Review. *Org. Prep. Proced. Int.* **1990**, *22*, 399–484.
(24) Karjalainen, O. K.; Koskinen, A. M. P. Diastereoselective Synthesis of Vicinal Amino Alcohols. *Org. Biomol. Chem.* **2012**, *10*, 4311–4326.
(25) Tsuchiya, S.; Cho, Y.; Konoki, K.; Nagasawa, K.; Oshima, Y.; Yotsu-Yamashita, M. Synthesis and Identification of Proposed Biosynthetic Intermediates of Saxitoxin in the Cyanobacterium Anabaena Circinalis (TA04) and the Dinoflagellate Alexandrium Tamarense (Axat-2). *Org. Biomol. Chem.* **2014**, *12*, 3016–3020.
(26) Astner, I.; Schulze, J. O.; van den Heuvel, J.; Jahn, D.; Schubert, W.-D.; Heinz, D. W. Crystal Structure of 5-Aminolevulinate Synthase, the First Enzyme of Heme Biosynthesis, and Its Link to XLSA in Humans. *EMBO J.* **2005**, *24*, 3166–3177.
(27) Mann, S.; Ploux, O. Pyridoxal-5′-Phosphate-Dependent Enzymes Involved in Biotin Biosynthesis: Structure, Reaction Mechanism and Inhibition. *Biochim. Biophys. Acta - Proteins Proteomics* **2011**, *1814*, 1459–1465.
(28) Gable, K.; Gupta, S. D.; Han, G.; Niranjanakumari, S.; Harmon, J. M.; Dunn, T. M. A Disease-Causing Mutation in the Active Site of Serine Palmitoyltransferase Causes Catalytic Promiscuity. *J. Biol. Chem.* **2010**, *285*, 22846–22852.
(29) Webster, S. P.; Alexeev, D.; Campopiano, D. J.; Watt, R. M.; Alexeeva, M.; Sawyer, L.; Baxter, R. L. Mechanism of 8-Amino-7-Oxononanoate Synthase: Spectroscopic, Kinetic, and Crystallographic Studies. *Biochemistry* **2000**, *39*, 516–528.
(30) Alexeev, D.; Alexeeva, M.; Baxter, R. L.; Campopiano, D. J.; Webster, S. P.; Sawyer, L. The Crystal Structure of 8-Amino-7-Oxononanoate Synthase: A Bacterial PLP-Dependent, Acyl-CoA-Condensing Enzyme. *J. Mol. Biol.* **1998**, *284*, 401–419.
(31) Manandhar, M.; Cronan, J. E. The Canonical Biotin Synthesis Enzyme, 8-Amino-7-Oxononanoate Synthase (BioF), Utilizes Different Acyl Chain Donors in *Bacillus Subtilis* and *Escherichia Coli*. *Appl. Environ. Microbiol.* **2017**, No. October, AEM.02084-17.
(32) Bojja, R. S.; Cerny, R. L.; Proctor, R. H.; Du, L. Determining the Biosynthetic Sequence in the Early Steps of the Fumonisin Pathway by Use of Three Gene-Disruption Mutants of Fusarium Verticillioides. *J. Agric. Food Chem.* **2004**, *52*, 2855–2860.
(33) Gerber, R.; Lou, L.; Du, L. A PLP-Dependent Polyketide Chain Releasing Mechanism in the Biosynthesis of Mycotoxin Fumonisins in Fusarium Verticillioides. *J. Am. Chem. Soc.* **2009**, *131*, 3148–3149.
(34) Dai, G. Z.; Han, W. B.; Mei, Y. N.; Xu, K.; Jiao, R. H.; Ge, H. M.; Tan, R. X. Pyridoxal-5'-Phosphate-Dependent Bifunctional Enzyme Catalyzed Biosynthesis of Indolizidine Alkaloids in Fungi. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 1174–1180.
(35) Schmidt, A.; Sivaraman, J.; Li, Y.; Larocque, R.; Barbosa, J. A. R. G.; Smith, C.; Matte,

A.; Schrag, J. D.; Cygler, M. Three-Dimensional Structure of 2-Amino-3-Ketobutyrate CoA Ligase from *Escherichia Coli* Complexed with a PLP−Substrate Intermediate: Inferred Reaction Mechanism. *Biochemistry* **2001**, *40*, 5151–5160.

(36) Stojanovski, B. M.; Hunter, G. A.; Na, I.; Uversky, V. N.; Jiang, R. H. Y.; Ferreira, G. C. 5-Aminolevulinate Synthase Catalysis: The Catcher in Heme Biosynthesis. *Mol. Genet. Metab.* **2019**, *128*, 178–189.

(37) Zhang, W.; Bolla, M. L.; Kahne, D.; Walsh, C. T. A Three Enzyme Pathway for 2-Amino-3-Hydroxycyclopent-2-Enone Formation and Incorporation in Natural Product Biosynthesis. *J. Am. Chem. Soc.* **2010**, *132*, 6402–6411.

(38) Petrícková, K.; Chronáková, A.; Zelenka, T.; Chrudimskỳ, T.; Pospíšil, S.; Petrícek, M.; Krištufek, V. Evolution of Cyclizing 5-Aminolevulinate Synthases in the Biosynthesis of Actinomycete Secondary Metabolites: Outcomes for Genetic Screening Techniques. *Front. Microbiol.* **2015**, *6*, 1–15.

(39) Williamson, N. R.; Simonsen, H. T.; Ahmed, R. A. A.; Goldet, G.; Slater, H.; Woodley, L.; Leeper, F. J.; Salmond, G. P. C. Biosynthesis of the Red Antibiotic, Prodigiosin, in Serratia: Identification of a Novel 2-Methyl-3-n-Amyl-Pyrroie (MAP) Assembly Pathway, Definition of the Terminal Condensing Enzyme, and Implications for Undecylprodigiosin Biosynthesis in Streptomyces. *Mol. Microbiol.* **2005**, *56*, 971–989.

(40) Williamson, N. R.; Fineran, P. C.; Leeper, F. J.; Salmond, G. P. C. The Biosynthesis and Regulation of Bacterial Prodiginines. *Nat. Rev. Microbiol.* **2006**, *4*, 887–899.

(41) Yard, B. A.; Carter, L. G.; Johnson, K. A.; Overton, I. M.; Dorward, M.; Liu, H.; McMahon, S. A.; Oke, M.; Puech, D.; Barton, G. J.; Naismith, J. H.; Campopiano, D. J. The Structure of Serine Palmitoyltransferase; Gateway to Sphingolipid Biosynthesis. *J. Mol. Biol.* **2007**, *370*, 870–886.

(42) Cerdeño, A. M.; Bibb, M. J.; Challis, G. L. Analysis of the Prodiginine Biosynthesis Gene Cluster of Streptomyces Coelicolor A3(2): New Mechanisms for Chain Initiation and Termination in Modular Multienzymes. *Chem. Biol.* **2001**, *8*, 817–829.

(43) Garneau-Tsodikova, S.; Dorrestein, P. C.; Kelleher, N. L.; Walsh, C. T. Protein Assembly Line Components in Prodigiosin Biosynthesis: Characterization of PigA,G,H,I,J. *J. Am. Chem. Soc.* **2006**, *128*, 12600–12601.

(44) Sakai-Kawada, F. E.; Ip, C. G.; Hagiwara, K. A.; Awaya, J. D. Biosynthesis and Bioactivity of Prodiginine Analogs in Marine Bacteria, Pseudoalteromonas: A Mini Review. *Front. Microbiol.* **2019**, *10*, 1–9.

(45) Rebets, Y.; Nadmid, S.; Paulus, C.; Dahlem, C.; Herrmann, J.; Hübner, H.; Rückert, C.; Kiemer, A. K.; Gmeiner, P.; Kalinowski, J.; Müller, R.; Luzhetskyy, A. Perquinolines A–C: Unprecedented Bacterial Tetrahydroisoquinolines Involving an Intriguing Biosynthesis. *Angew. Chem. Int. Ed.* **2019**, *58*, 12930–12934.

(46) Kawata, J.; Naoe, T.; Ogasawara, Y.; Dairi, T. Biosynthesis of the Carbonylmethylene Structure Found in the Ketomemicin Class of Pseudotripeptides. *Angew. Chem. Int. Ed.* **2017**, *56*, 2026–2029.

(47) Masuo, S.; Tsuda, Y.; Namai, T.; Minakawa, H.; Shigemoto, R.; Takaya, N. Enzymatic Cascade in Pseudomonas That Produces Pyrazine from α-Amino Acids. *ChemBioChem* **2020**, *21*, 353–359.

(48) Jahan, N.; Potter, J. A.; Sheikh, M. A.; Botting, C. H.; Shirran, S. L.; Westwood, N. J.; Taylor, G. L. Insights into the Biosynthesis of the Vibrio Cholerae Major Autoinducer CAI-1 from the Crystal Structure of the PLP-Dependent Enzyme CqsA. *J. Mol. Biol.* **2009**, *392*,

763–773.
(49) Wei, Y.; Perez, L. J.; Ng, W.-L.; Semmelhack, M. F.; Bassler, B. L. Mechanism of Vibrio Cholerae Autoinducer-1 Biosynthesis. *ACS Chem. Biol.* **2011**, *6*, 356–365.
(50) Spirig, T.; Tiaden, A.; Kiefer, P.; Buchrieser, C.; Vorholt, J. A.; Hilbi, H. The Legionella Autoinducer Synthase LqsA Produces an α-Hydroxyketone Signaling Molecule. *J. Biol. Chem.* **2008**, *283*, 18113–18123.
(51) Haack, F. S.; Poehlein, A.; Kröger, C.; Voigt, C. A.; Piepenbring, M.; Bode, H. B.; Daniel, R.; Schäfer, W.; Streit, W. R. Molecular Keys to the Janthinobacterium and Duganella Spp. Interaction with the Plant Pathogen Fusarium Graminearum. *Front. Microbiol.* **2016**, *7*, 1668.
(52) Bordwell, F. G.; Boutan, P. J. Conjugative Effects in Divalent Sulfur Groupings. *J. Am. Chem. Soc.* **1956**, *78*, 854–860.
(53) Padmakumar, R.; Gantla, S.; Banerjee, R. A Rapid Method for the Synthesis of Methylmalonyl-Coenzyme A and Other CoA-Esters. *Anal. Biochem.* **1993**, *214*, 318–320.
(54) Worthington, A. S.; Burkart, M. D. One-Pot Chemo-Enzymatic Synthesis of Reporter-Modified Proteins. *Org. Biomol. Chem.* **2006**, *4*, 44–46.
(55) Sieber, S. A.; Tao, J.; Walsh, C. T.; Marahiel, M. A. Peptidyl Thiophenols as Substrates for Nonribosomal Peptide Cyclases. *Angew. Chem. Int. Ed.* **2004**, *43*, 493–498.
(56) Hansen, D. A.; Koch, A. A.; Sherman, D. H. Substrate Controlled Divergence in Polyketide Synthase Catalysis. *J. Am. Chem. Soc.* **2015**, *137*, 3735–3738.
(57) Hansen, D. A.; Rath, C. M.; Eisman, E. B.; Narayan, A. R. H.; Kittendorf, J. D.; Mortison, J. D.; Yoon, Y. J.; Sherman, D. H. Biocatalytic Synthesis of Pikromycin, Methymycin, Neomethymycin, Novamethymycin, and Ketomethymycin. *J. Am. Chem. Soc.* **2013**, *135*, 11232–11238.
(58) Shoolingin-Jordan, P. M.; Al-Daihan, S.; Alexeev, D.; Baxter, R. L.; Bottomley, S. S.; Kahari, I. D.; Roy, I.; Sarwar, M.; Sawyer, L.; Wang, S. F. 5-Aminolevulinic Acid Synthase: Mechanism, Mutations and Medicine. *Biochim. Biophys. Acta - Proteins Proteomics* **2003**, *1647*, 361–366.
(59) Toney, M. D. Controlling Reaction Specificity in Pyridoxal Phosphate Enzymes. *Biochim. Biophys. Acta - Proteins Proteomics* **2011**, *1814*, 1407–1418.
(60) Abboud, M. M.; Jordan, P. M.; Akhtar, M. Biosynthesis of 5-Aminolevulinic Acid: Involvement of a Retention–Inversion Mechanism. *J. Chem. Soc., Chem. Commun.* **1974**, *0*, 643–644.
(61) Bashir, Q.; Rashid, N.; Akhtar, M. Mechanism and Substrate Stereochemistry of 2-Amino-3-Oxobutyrate CoA Ligase: Implications for 5-Aminolevulinate Synthase and Related Enzymes. *Chem. Commun.* **2006**, *0*, 5065–5067.
(62) Bhor, V. M.; Dev, S.; Vasanthakumar, G. R.; Kumar, P.; Sinha, S.; Surolia, A. Broad Substrate Stereospecificity of the Mycobacterium Tuberculosis 7-Keto-8-Aminopelargonic Acid Synthase: Spectroscopic and Kinetic Studies. *J. Biol. Chem.* **2006**, *281*, 25076–25088.
(63) Mann, S.; Colliandre, L.; Labesse, G.; Ploux, O. Inhibition of 7,8-Diaminopelargonic Acid Aminotransferase from Mycobacterium Tuberculosis by Chiral and Achiral Anologs of Its Substrate: Biological Implications. *Biochimie* **2009**, *91*, 826–834.
(64) Hanada, K. Serine Palmitoyltransferase, a Key Enzyme of Sphingolipid Metabolism. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **2003**, *1632*, 16–30.
(65) Lukowski, A. L.; Mallik, L.; Hinze, M. E.; Carlson, B. M.; Ellinwood, D. C.; Pyser, J. B.; Koutmos, M.; Narayan, A. R. H. Substrate Promiscuity of a Paralytic Shellfish Toxin

Amidinotransferase. *ACS Chem. Biol.* **2020**, *15*, 626–631.
(66) Mihali, T. K.; Kellmann, R.; Neilan, B. A. Characterisation of the Paralytic Shellfish Toxin Biosynthesis Gene Clusters in Anabaena Circinalis AWQC131C and Aphanizomenon Sp. NH-5. *BMC Biochem.* **2009**, *10*, 8.
(67) Mihali, T. K.; Carmichael, W. W.; Neilan, B. A. A Putative Gene Cluster from a Lyngbya Wollei Bloom That Encodes Paralytic Shellfish Toxin Biosynthesis. *PLoS ONE* **2011**, *6*, e14657.
(68) Sendall, B. C.; McGregor, G. B. Cryptic Diversity within the Scytonema Complex: Characterization of the Paralytic Shellfish Toxin Producer Heterosyctonema Crispum, and the Establishment of the Family Heteroscytonemataceae (Cyanobacteria/Nostocales). *Harmful Algae* **2018**, *80*, 158–170.
(69) Diefenbach, X. W.; Farasat, I.; Guetschow, E. D.; Welch, C. J.; Kennedy, R. T.; Sun, S.; Moore, J. C. Enabling Biocatalysis by High-Throughput Protein Engineering Using Droplet Microfluidics Coupled to Mass Spectrometry. *ACS Omega* **2018**, *3*, 1498–1508.
(70) Ngernsutivorakul, T.; Steyer, D. J.; Valenta, A. C.; Kennedy, R. T. In Vivo Chemical Monitoring at High Spatiotemporal Resolution Using Microfabricated Sampling Probes and Droplet-Based Microfluidics Coupled to Mass Spectrometry. *Anal. Chem.* **2018**, *90*, 10943–10950.
(71) Miyazaki, K. *MEGAWHOP Cloning: A Method of Creating Random Mutagenesis Libraries via Megaprimer PCR of Whole Plasmids*, 1st ed.; Elsevier Inc., 2011; Vol. 498.
(72) Steyer, D. J.; Kennedy, R. T. High-Throughput Nanoelectrospray Ionization-Mass Spectrometry Analysis of Microfluidic Droplet Samples. *Anal. Chem.* **2019**.
(73) Holland-Moritz, D. A.; Wismer, M. K.; Mann, B. F.; Farasat, I.; Devine, P.; Guetschow, E. D.; Mangion, I.; Welch, C. J.; Moore, J. C.; Sun, S.; Kennedy, R. T. Mass Activated Droplet Sorting (MADS) Enables High-Throughput Screening of Enzymatic Reactions at Nanoliter Scale. *Angew. Chem. Int. Ed.* **2020**, *59*, 4470–4477.
(74) Lin, H.-S.; Paquette, L. a. A Convenient Method for Determining the Concentration of Grignard Reagents. *Synth. Commun.* **1994**, *24*, 2503–2506.
(75) Wang, L.; Qiao, J.; Wei, J.; Liang, Z.; Xu, X.; Li, N. Air-Stable Binuclear Titanium(IV) Salophen Perfluorobutanesulfonate with Zinc Power Catalytic System and Its Application to C–S and C–Se Bond Formation. *Tetrahedron* **2020**, *76*, 130750.
(76) Kwon, Y.-D.; La, M. T.; Kim, H.-K. Aerobic Oxidative Esterification and Thioesterification of Aldehydes Using Dibromoisocyanuric Acid under Mild Conditions: No Metal Catalysts Required. *New J. Chem.* **2018**, *42*, 10833–10841.
(77) Böttcher, T.; Sieber, S. A. β-Lactones as Privileged Structures for the Active-Site Labeling of Versatile Bacterial Enzyme Classes. *Angew. Chem. Int. Ed.* **2008**, *47*, 4600–4603.
(78) Frei, R.; Wodrich, M. D.; Hari, D. P.; Borin, P. A.; Chauvier, C.; Waser, J. Fast and Highly Chemoselective Alkynylation of Thiols with Hypervalent Iodine Reagents Enabled through a Low Energy Barrier Concerted Mechanism. *J. Am. Chem. Soc.* **2014**, *136*, 16563–16573.
(79) Bogonda, G.; Patil, D. V.; Kim, H. Y.; Oh, K. Visible-Light-Promoted Thiyl Radical Generation from Sodium Sulfinates: A Radical–Radical Coupling to Thioesters. *Org. Lett.* **2019**, *21*, 3774–3779.
(80) Kim, S.; Kim, S.; Otsuka, N.; Ryu, I. Tin-Free Radical Carbonylation: Thiol Ester Synthesis Using Alkyl Allyl Sulfone Precursors, Phenyl Benzenethiosulfonate, and CO. *Angew. Chem. Int. Ed.* **2005**, *44*, 6183–6186.
(81) Karier, P.; Ungeheuer, F.; Ahlers, A.; Anderl, F.; Wille, C.; Fürstner, A. Metathesis at an

Implausible Site: A Formal Total Synthesis of Rhizoxin D. *Angew. Chem. Int. Ed.* **2019**, *58*, 248–253.
(82) Guan, X.; Drake, M. R.; Tan, Z. Total Synthesis of Human Galanin-Like Peptide through an Aspartic Acid Ligation. *Org. Lett.* **2013**, *15*, 6128–6131.
(83) Wang, F.; Wang, Y.; Ji, J.; Zhou, Z.; Yu, J.; Zhu, H.; Su, Z.; Zhang, L.; Zheng, J. Structural and Functional Analysis of the Loading Acyltransferase from Avermectin Modular Polyketide Synthase. *ACS Chem. Biol.* **2015**, *10*, 1017–1025.
(84) Boudreau, P. D.; Miller, B. W.; McCall, L. I.; Almaliti, J.; Reher, R.; Hirata, K.; Le, T.; Siqueira-Neto, J. L.; Hook, V.; Gerwick, W. H. Design of Gallinamide A Analogs as Potent Inhibitors of the Cysteine Proteases Human Cathepsin L and Trypanosoma Cruzi Cruzain. *J. Med. Chem.* **2019**, *62*, 9026–9044.
(85) Barroso, R.; Escribano, M.; Cabal, M. P.; Valdés, C. Tosylhydrazide-Promoted Diastereoselective Intramolecular 1,3-Dipolar Cycloadditions: Synthesis of Tetrahydropyrrolo[3,4-c]Pyrazoles. *European J. Org. Chem.* **2014**, *2014*, 1672–1683.
(86) Tsunematsu, Y.; Nishimura, S.; Hattori, A.; Oishi, S.; Fujii, N.; Kakeya, H. Isolation, Structure Elucidation, and Total Synthesis of Tryptopeptins A and B, New TGF-β Signaling Modulators from Streptomyces Sp. *Org. Lett.* **2015**, *17*, 258–261.
(87) Pfeifer, B. A.; Admiraal, S. J.; Gramajo, H.; Cane, D. E.; Khosla, C. Biosynthesis of Complex Polyketides in a Metabolically Engineered Stain of E. Coli. *Science* **2001**, *291*, 1790–1792.
(88) Eschenfeldt, W. H.; Lucy, S.; Millard, C. S.; Joachimiak, A.; Mark, I. D. A Family of LIC Vectors for High-Throughput Cloning and Purification of Proteins. In *Methods in Molecular Biology*; Doyle, S. A., Ed.; Humana Press: Totowa, NJ, 2009; Vol. 498, pp 105–115.
(89) Dorrestein, P. C.; Bumpus, S. B.; Calderone, C. T.; Garneau-Tsodikova, S.; Aron, Z. D.; Straight, P. D.; Kolter, R.; Walsh, C. T.; Kelleher, N. L. Facile Detection of Acyl and Peptidyl Intermediates on Thiotemplate Carrier Domains via Phosphopantetheinyl Elimination Reactions during Tandem Mass Spectrometry. *Biochemistry* **2006**, *45*, 12756–12766.

Chapter 4: SxtA AONS-Mediated Deuterium Labeling

Reprinted (adapted) with permission from "Biocatalytic, Stereoselective Deuteration of α-Amino Acids and Methyl Esters". Chun, S. W.; Narayan, A. R. H. *ACS Catal.* **2020**, *10,* 7413-7418.

Summary

α-^{2}H amino acids are valuable precursors toward labeled pharmaceutical agents and tools for studying biological systems; however, these molecules are costly to purchase and challenging to synthesize in a site- and stereoselective manner. Here, we show that an α-oxoamine synthase that evolved for saxitoxin biosynthesis, SxtA AONS, can produce a range of α-^{2}H amino acids and esters site- and stereoselectively using D_2O as the deuterium source. Additionally, we demonstrate the utility of this operationally simple reaction on preparative scale in the stereoselective chemoenzymatic synthesis of a deuterated analog of safinamide, a drug used to treat Parkinson's disease. Having established that SxtA AONS can generate its nucleophilic quinonoid intermediate in the absence of a thioester partner, we will also be able to harness this biocatalyst for the formation of other amino acid derivatives.

4.1 Introduction

During our studies in the SxtA AONS-mediated synthesis of non-native α-amino ketones, we also considered whether this enzyme could elaborate α-amino acids into α-tetra-substituted derivatives (see **4.4**, Figure 4.1). These products would arise from the key nucleophilic quinonoid

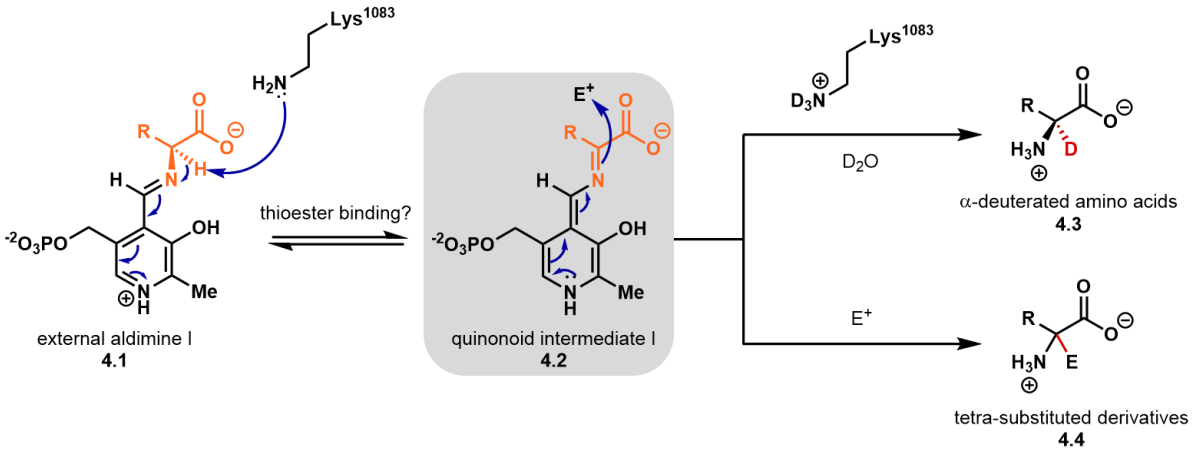


Figure 4.1. SxtA AONS-catalyzed diversification of α-amino acids through quinonoid intermediate I.

intermediate **4.2** adding to other classes of electrophiles rather than thioesters. Whether some other factor is required to trigger deprotonation of the external aldimine intermediate **4.1** in SxtA AONS to generate **4.2** was unknown.

For numerous members of the α-oxoamine synthase (AOS) family, deprotonation does not proceed until the thioester substrate was added to the reaction solution.[1–4] In reactions with SxtA AONS, the presence of compatible thioesters would result in mixture of competing α-amino ketone and tetra-substituted products. However, there are reports of three homologs, *Bacillus sphaericus* BioF and aminolevulinic acid synthase (ALAS) from *Rhodobacter sphaeroides* and *Mus musculus*,[5–7] forming quinonoid I with only amino acid present. Other BioF[2] and ALAS[4] homologs also belong to the "thioester required" group, and this property also is not apparently predicted by the primary amino acid sequence alone.

In the mid and late 1900s, it was debated whether the step following amino acid-PLP imine formation is deprotonation or decarboxylation. The existence of the non-decarboxylating 2-amino-3-ketobutyric acid CoA ligase (KBL, see Chapter 3.1) in the AOS protein family, and multiple

isotope labeling experiments summarized in Figure 4.2 strongly support that the step following amino acid binding in the AOS mechanism is indeed deprotonation.[2,5,6,8,9] Figure 4.2B depicts the

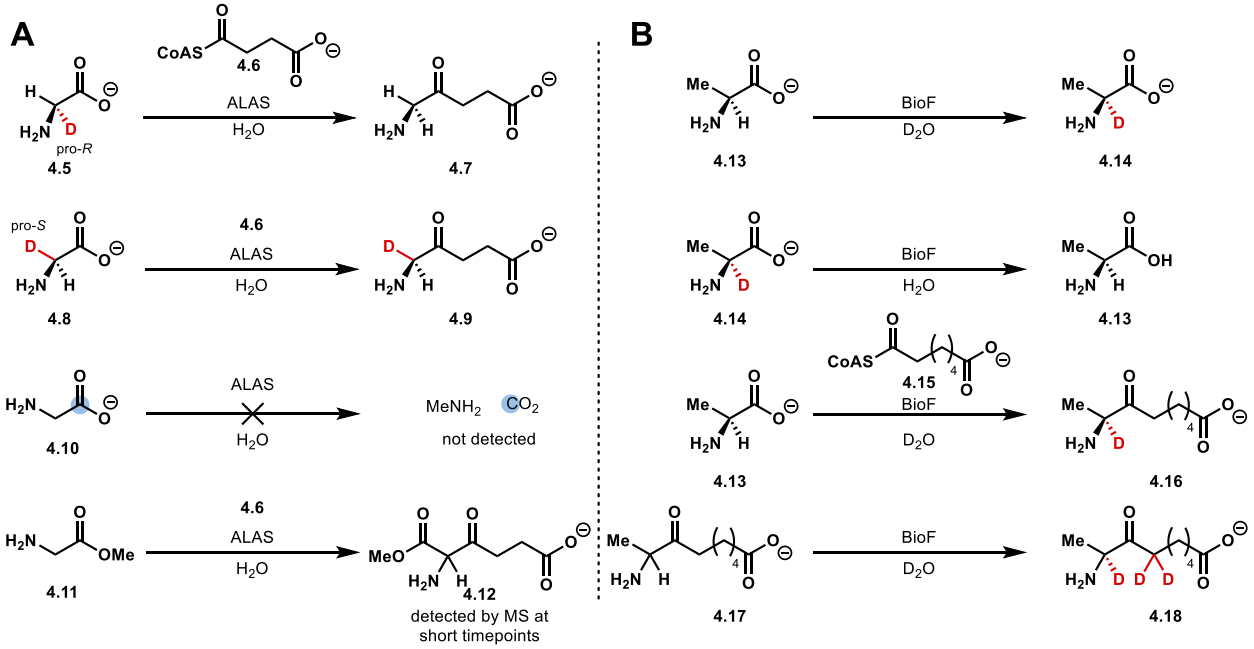


Figure 4.2. Isotopic labeling experiments to determine the mechanism of AOS enzymes. (A) Reactions with HemA or aminolevulinic acid synthase (ALAS). (B) reactions with BioF.

results of the initial experiments aimed at ascertaining to which group SxtA AONS belongs, "thioester required", or "thioester-less" deprotonation. *Microseira wollei* SxtA AONS and L-Arg (**4.19**) were incubated overnight in D_2O-exchanged buffer and analyzed by MS (Figure 4.3). Approximately 18% of the arginine population now showed a mass of 1 Da higher than the substrate, indicating the replacement of one non-exchangeable hydrogen with a deuterium atom. This placed *M. wollei* SxtA AONS (as well as the *C. raciborskii* AONS homolog in our possession) in the "thioester-less" group, which would allow us to pursue the formation of tetra-substituted amino acid derivatives (see **4.4**, Figure 4.1). This result prompted the exploration of SxtA AONS-mediated α-deuteration as a synthetically useful reaction.

Deuterated compounds are important tools in fundamental research and commercial endeavors including drug discovery and development.[10–13] Deuterated drugs are often metabolized more

slowly than the analogous protio-compounds due to the kinetic isotope effect, affording longer half-lives and allowing for lower doses.[14] The first deuterated drug candidate fludalanine (**4.22**,

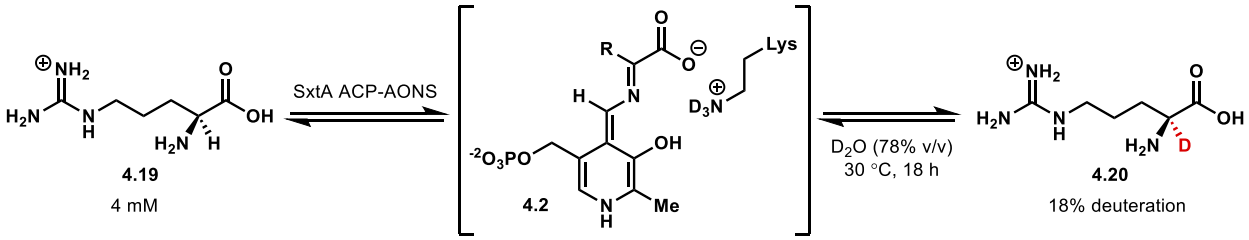


Figure 4.3. α-deuteration of L-Arg by SxtA AONS in deuterated buffer in the absence of a thioester partner.

see Figure 4.4) was initially tested in the late 1970s as a mixture with cycloserine for inhibition of bacterial alanine racemase.[15] For clinical trials, **4.22** was synthesized on gram-scale in four synthetic steps and one chiral resolution step.[16] The deuterated analog was metabolized 2-3 times more slowly than the protio-compound, but eventually failed trials due to buildup of a toxic downstream metabolite.[10,17] Scientists have continued developing deuterated bioactive compounds, with the FDA also recognizing the potential benefits of deuterated pharmaceutical agents. The first deuterated analog of an existing drug, deutetrabenazine (**4.23**), was approved in 2017 for its doubled half-life over tetrabenazine.[18]

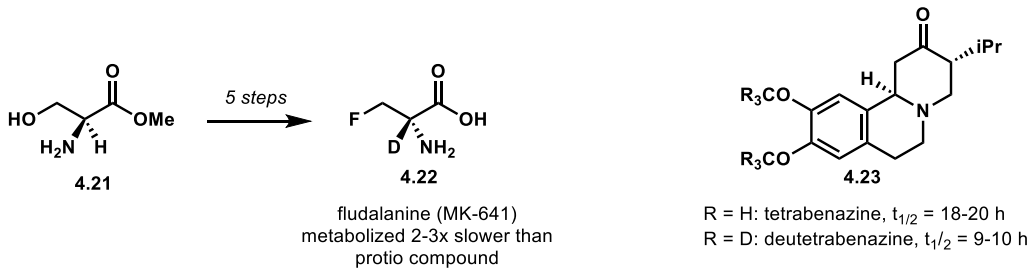


Figure 4.4. Historical pharmaceutical candidates and products containing deuterium.

The value of deuterated compounds creates a demand for methods to selectively introduce isotopic labels into available building blocks. For example, α-^{2}H amino acids are potential pharmaceutical precursors,[19] can be incorporated into proteins to improve NMR signal,[20] and are

tools to probe protein and natural product metabolism.[21,22] Despite the value of these molecules, synthetic access is challenging, with difficulties in achieving site- and stereoselectivity in the

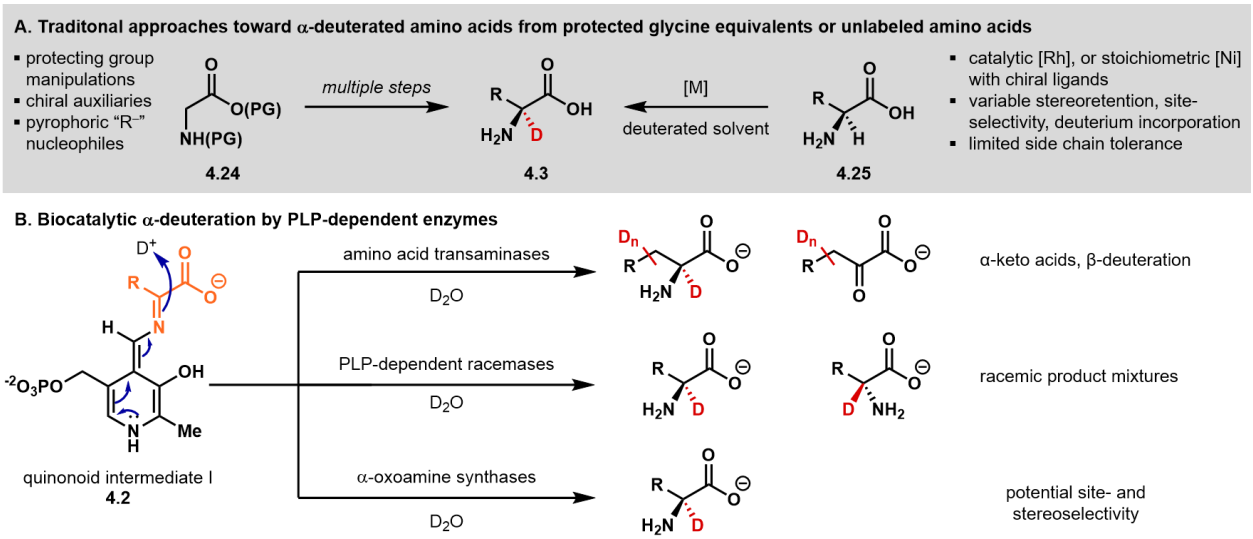


Figure 4.5. Preparations of deuterated compounds and amino acids.
(A) Classical methods for synthesizing α-deuterated amino acids are lengthy (left). Recent chemical methods starting from the corresponding protio-amino acids vary in activity (right). (B) Potential biocatalytic deuteration by three classes of PLP-dependent enzymes.

deuterium incorporation step. Established methods toward α-^{2}H amino acids follow multi-step routes where protecting groups and chiral auxiliaries are required (see **4.24** to **4.3**, Figure 4.5A), accounting for the hundred-fold higher cost for enantioenriched α-^{2}H-amino acids relative to their protio- counterparts.[23–25] Recent approaches offer more efficient access by targeting deuterated compounds directly from the corresponding unlabeled acids (see **4.25** to **4.3**, Figure 4.5A). Although these methods provide high-levels of deuterium incorporation, they often fall short in achieving site- and stereoselectivity on a wide array of substrates.[26,27]

A biocatalytic approach to the transformation of protio-amino acids to α-deuterated analogs offers several potential advantages, including precise stereoselectivity, the elimination of protecting groups, and mild, sustainable reaction conditions.[28] Deuterium sources such as D_2 and D_2O are economical and available in bulk.[29] Any of the pyridoxal phosphate (PLP)-dependent

subclasses that begin with the deprotonation of the α-carbon could potentially be a biocatalyst for deuterium incorporation.[30,31] In contrast to PLP-dependent transaminases and racemases, which catalyze deuterated amino acid formation with some substrate flexibility but rarely have control over site- and stereoselectivity (Figure 4.5B).[32,33] α-oxoamine synthases (AOS) have been shown to install a deuterium atom at the α-position of select amino acids stereoselectively. For example, in mechanistic studies of AOS proteins BioF and serine palmitoyltransferase (SPT), incubation in D_2O led to α-deuteration of their native substrates L-Ala and L-Ser, respectively.[3,5] The deuterated products retained the L- configuration.

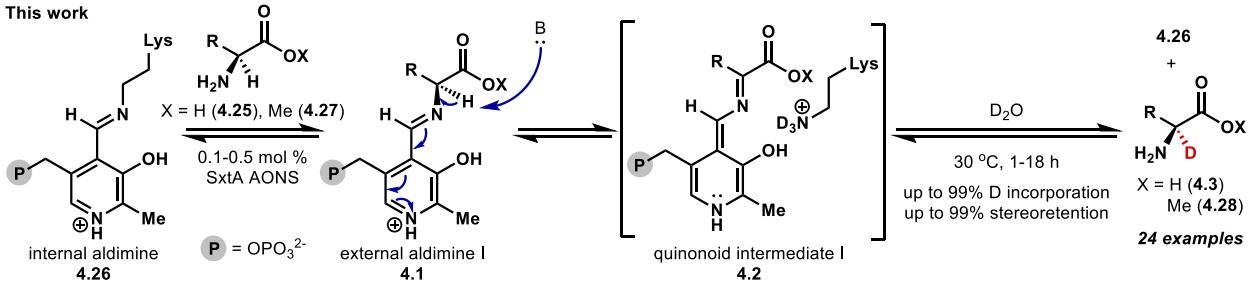


Figure 4.6. SxtA AONS mediates α-deuterium incorporation on a wide range of substrates.

After demonstrating that SxtA AONS accepts multiple structurally unrelated α-amino acid substrates in addition to its native L-arginine for conversion into α-amino ketones (see Chapter 3.3), we reasoned that those seven non-native acids and perhaps others would also be substrates for α-deuteration. Here, we show that SxtA AONS is indeed an effective catalyst for α-deuteration of select L-amino acids and a broad panel of α-amino ester substrates (Figure 4.6). These deuterated building blocks can be generated on preparative-scale and employed in multi-step synthesis as we have demonstrated in the synthesis of a deuterated form of the Parkinson's drug, safinamide.[34,35]

4.2 Deuterium Incorporation into Arg-Based Compounds

Of the two methods to detect the presence of a quinonoid intermediate, UV-vis spectroscopy or isotopic labeling, we selected the latter as for its relative ease of interpretation. The spectroscopic details of SxtA AONS are discussed in Chapter 5.3. Our activity test consisted of incubating enzyme and an amino acid in deuterated buffer followed by MS analysis.

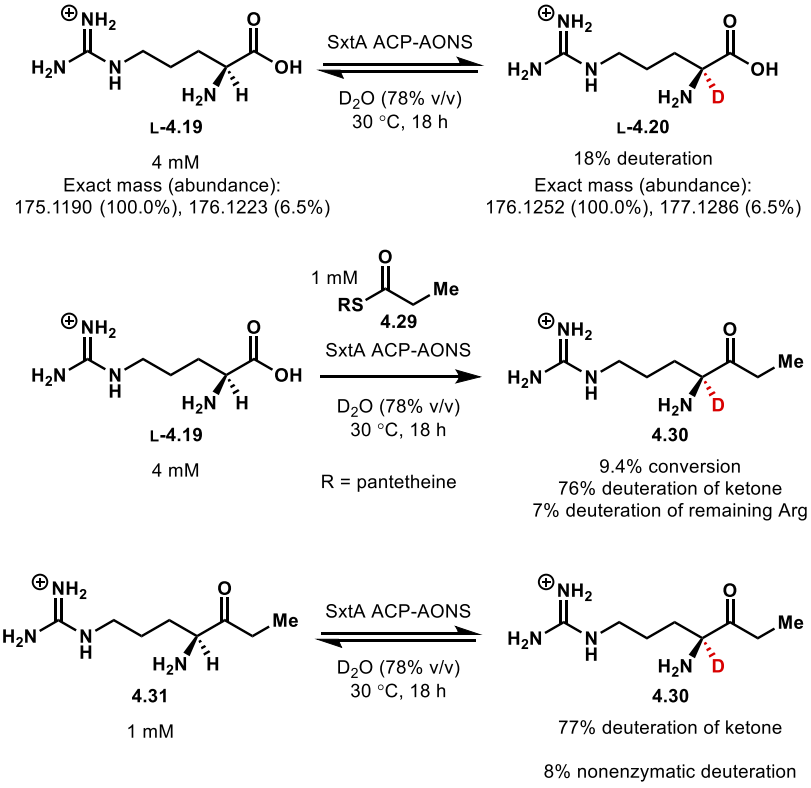


Figure 4.7. Initial experiments of SxtA ACP-AONS in D_2O.
SxtA AONS incorporates deuterium into the native amino acid substrate L-Arg (**4.19**) at low levels, and at near theoretical maximum levels of the native ketone product (**4.31**).

Several factors were considered prior to and during the analysis. Enzymatic reactions were quenched with a protio rather than deuterated (i.e., CH_3OH instead of CD_3OD) organic solvent to ensure that all exchangeable hydrogens were protons, and that any change in the exact mass was due to the formation of a new C–D bond, not heteroatom–D. Additionally, all ordinary protio-compound samples will contain isotopologs of the monoisotopic mass due to the presence of naturally occurring heavy isotopes such as ^{13}C, ^{15}N, and ^{34}S. For example, in arginine, an

isotopolog with one ^{13}C atom (176.1223 Da) is predicted to occur at 6.5% abundance than that of the monoisotopic mass (175.1190 Da). The difference in exact mass between a ^{13}C and ^{2}H isotopolog (176.1252 Da) is 0.0029 Da, too small to be resolved on our TOF and Q-TOF MS instruments. Therefore, these calculated abundances of heavy isotopologs were subtracted from the isotopic distributions to correct for naturally occurring heavy isotopes and compared to no-enzyme negative controls in D_2O and H_2O.

Initial experiments with *M. wollei* SxtA ACP-AONS and 4 mM L-Arg (**L-4.19**) incubated in deuterated buffer (78:22 v/v D_2O to H_2O) showed 18% deuterium incorporation after overnight incubation, equivalent to at least 36 deuteration turnovers (Figure 4.7). A positive control condensation reaction also containing propionyl-pantetheine (**4.29**) in addition to enzyme and L-Arg resulted in 94 μM (9.4% conversion or <5 turnovers) of the native ketone product **4.31**. However, 76% of the ketone formed had incorporation of at least one deuterium atom (**4.30**). In a separate reaction containing enzyme and 1 mM of the synthetic standard of **4.31**, the total deuterium incorporation into the ketone was also 77%. This is equivalent to at least 154 enzyme deuteration turnovers, indicating that SxtA AONS may be more active in deprotonation and protonation than condensation. When the volumetric fraction of D_2O was raised to >99:1 by preparing the buffers directly in D_2O and pre-equilibrating the enzyme, the deuterium incorporation in 4 mM L-Arg increased from 18% to 37% (about 74 turnovers).

Taken together, these data indicate that *M. wollei* SxtA AONS falls in the "thioester-less deprotonation" group of AOS enzymes. Because the protio-substrates and the deuterated products are in equilibrium with each other, the theoretical maximum deuterium incorporation is approximately equal to the D_2O:H_2O ratio. Any possible equilibrium isotope effects were not measured.[36]

The SxtA AONS product ketone **4.31** is also a better compound for AONS-mediated deuterium incorporation than the native amino acid substrate, L-Arg (**L-4.19**). One contributing factor may be the pK_a of the alpha carbon, which is several units lower in an α-amino ketone than an α-amino acid.[5,37] In a negative control reaction lacking enzyme, 78% D_2O conditions tested above afforded 8% of ketone **4.30** (at a starting concentration of 1 mM) with deuterium incorporation, while no nonenzymatic deuterium incorporation into L-Arg was detected. A second consideration may be that the binding of the ketone better induces a protein conformational change that favors deprotonation and formation of quinonoid intermediates.[2,37]

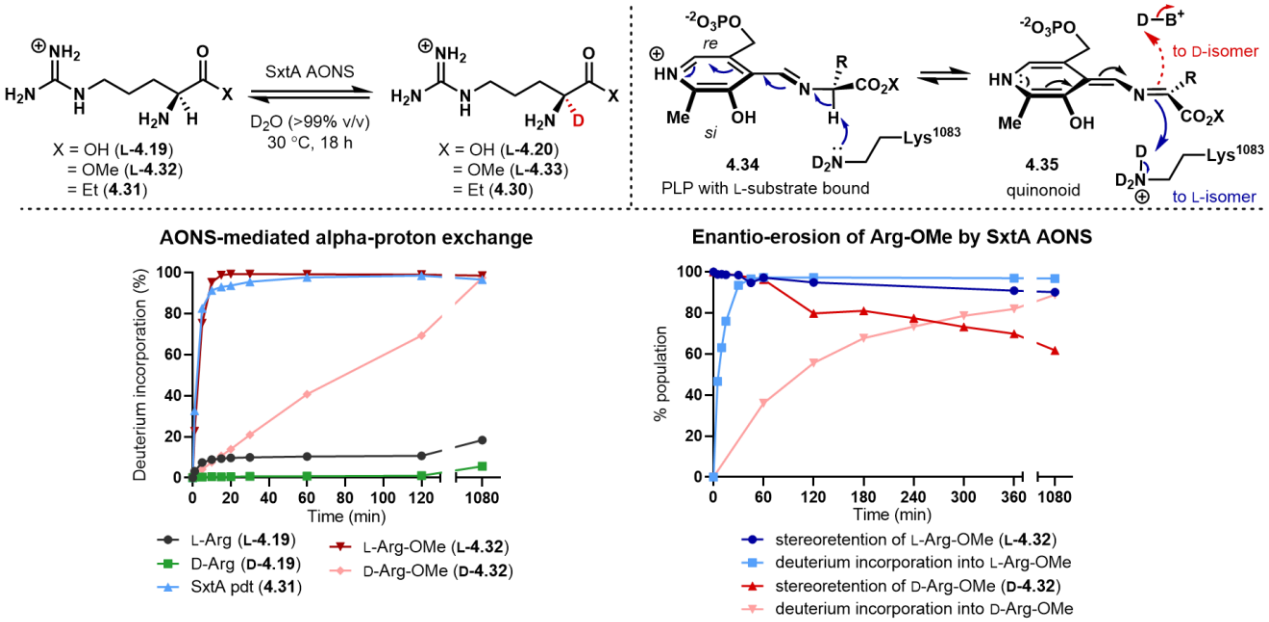


Figure 4.8. Deuteration of arginine-related substrates.
(A) α-deuterated compound formation. (B) Potential mechanisms of proton/deuterium transfers in SxtA AONS. (C) Timecourse of α-deuteration by mass spectrometry. (D) Comparison of deuterium incorporation and enantiomeric composition when starting with L-Arg-OMe (blue) or D-Arg-OMe (red).

After discovering that the standalone domain SxtA AONS is sufficient to accomplish deuterium incorporation, we conducted all further experiments with SxtA AONS rather than the SxtA ACP-AONS didomain. To compare the approximate rates of deuterium incorporation in

these arginine-based compounds, we conducted time course experiments with both arginine (**4.19**) and arginine methyl ester enantiomers[38] (**4.32**, all at 4 mM) and the product standard (**4.31**, at 2 mM) in >99% D_2O buffer. Under these conditions, approximately 20% of the L-acid had one deuterium atom incorporated as determined by LC-MS analysis (Figure 4.8C, black line). Whereas deuterium labeling was moderate for L-Arg, both ketone **4.31** and its structural mimic L-Arg-OMe reached maximum theoretical deuteration within 15 minutes (100 and 200 turnovers, respectively). Interestingly, substrates deuterated at a low level by SxtA AONS were readily transformed into compounds that were efficiently deuterated through esterification to the corresponding methyl ester (see D-Arg vs. D-Arg-OMe). Minor deuterium incorporation into the acids and esters was observed in controls lacking active enzyme, while the lower pK_a of ketone **4.31** led to 15% nonenzymatic deuteration.[5]

The site of deuterium incorporation was confirmed to be the α-carbon of L-Arg-OMe (**L-4.32**). In a phosphate-buffered reaction monitored by ^1H NMR, the α-proton triplet signal at 4.09 ppm disappeared over 45 min, demonstrating complete replacement by deuterium (Figure 4.9). No decrease in the beta proton signals was observed, even at extended timepoints indicating a higher level of site-selectivity for SxtA AONS compared with deuteration mediated by other PLP-dependent enzymes.[32,39,40] The exception to this site-selectivity was observed only with ketone **4.31**, which possesses additional acidic protons alpha to the carbonyl group.[5] No more than a single deuterium atom was incorporated into non-glycine-based amino acid and methyl ester-based substrates.

As SxtA AONS deuterates both enantiomers of Arg-OMe, we next investigated the configuration of the deuterated products (**4.33**). In prior studies of AOS proteins, the catalytic lysine residue (Lys1083 in *M. wollei* SxtA) is proposed to stereospecifically mediate both the

initial deprotonation of the L-amino acid substrate and the subsequent reprotonation of the α-position to generate L- products (Figure 4.8B, blue arrows).[41,42] Notably, this residue is located on the *si* face of PLP in all published AOS crystal structures.[43–47] On the basis of this general mechanism, the deprotonation and deuteration of D-Arg-OMe (**D-32**) is unexpected, but consistent with our previous observation that D-Arg is also converted to ketone **4.31** at low levels with unknown stereochemistry (see Chapter 3.2). Ongoing structural studies in collaboration with Prof. Janet Smith's group aim to distinguish if SxtA AONS accommodates different binding modes[48] for each Arg-OMe enantiomer or if an alternative active site residue or water can serve as a base[49] to account for activity on D-configured substrates (Figure 4.8B, red arrows, and Figure 4.S29).

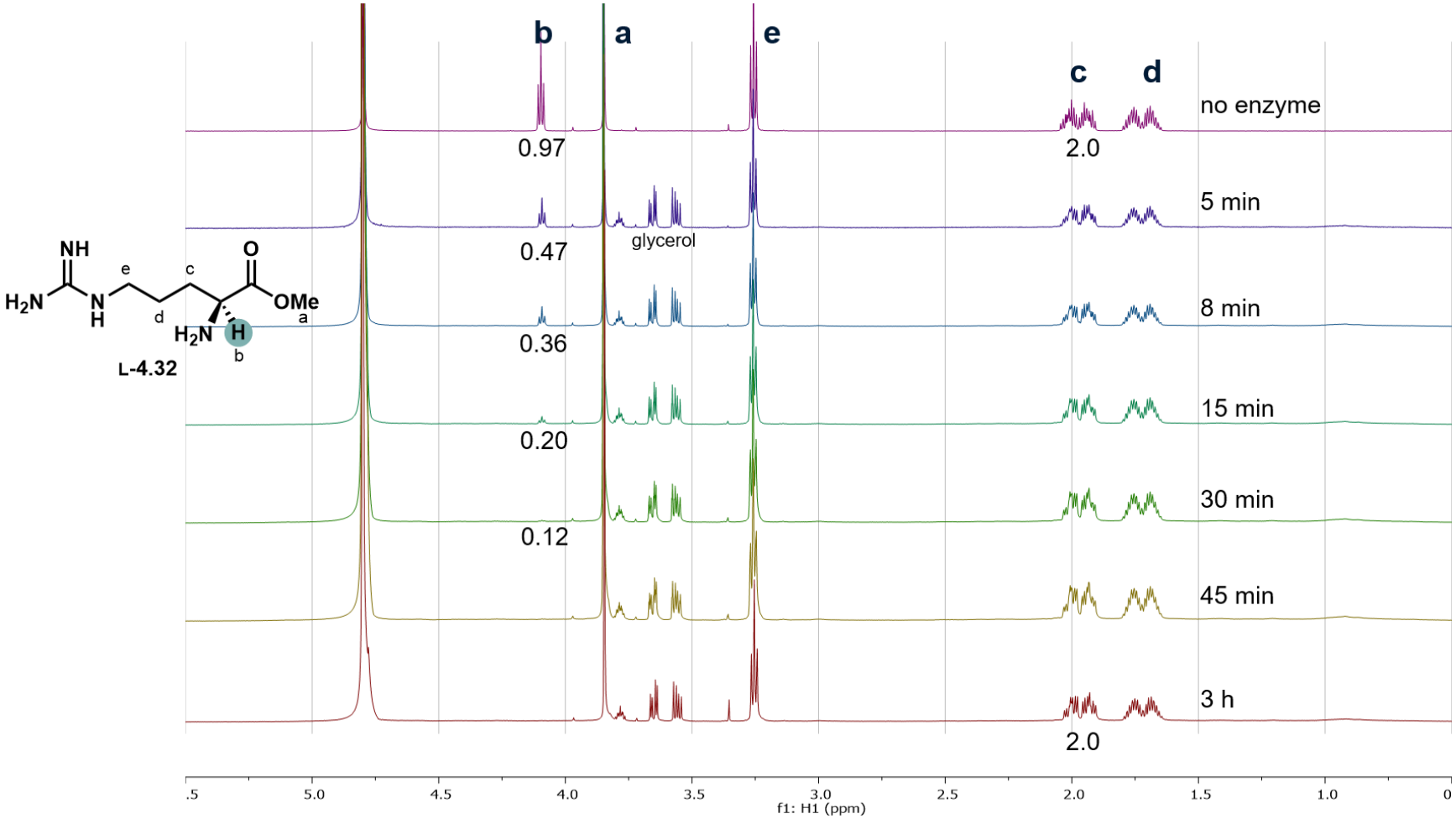


Figure 4.9. ^{1}H NMR of L-Arg-OMe incubated with AONS over 3 h, confirming exchange of the α-proton (labeled b) to deuterium. Peaks present between 3.7-3.5 ppm are attributed to glycerol from the enzyme storage buffer. The integration of the beta protons (labeled c) remained at 2.0 throughout the experiment.

To determine the degree of stereocontrol exhibited by SxtA AONS, 20 mM of each Arg-OMe enantiomer was separately incubated with 20 μM SxtA AONS (1000 turnovers for full conversion). The resulting material was assessed for deuterium labeling and enantiopurity at various timepoints. *In situ* Fmoc protection after quenching the enzymatic reaction conveniently lowered the polarity of the α-amino esters and provided a UV-active chromophore that could be monitored. Analysis of Fmoc-derivatized compounds by chiral supercritical fluid chromatography (SFC) provided the enantiopurity of each product (see Table 4.S6). For L-Arg-OMe, complete deuterium incorporation and a high degree of retention of the L- configuration was observed at 45 min, with slight enantio-erosion to 90:10 L to D occurring by 18 h (Figure 4.8D, blue lines). At this higher substrate concentration, D-Arg-OMe reached 89% deuteration (red lines, 890 turnovers). The rate of enantio-erosion was faster than that of the L-enantiomer, ending at a 38:62 L to D enantiomeric ratio after 18 h. For both enantiomers of ester **4.32**, the retention of the original configuration does not track directly with the deuteration level, nor is the reprotonation step observed to be completely stereoselective. The reprotonation step for L-Arg-OMe is generally stereoretentive to afford primarily the deuterated L-isomer. The putative reprotonation step with a D-Arg-OMe substrate is less selective, producing significant quantities of L-product. The mechanism of deprotonation and reprotonation remains unclear in the absence of a SxtA AONS crystal structure.

4.3 Deuterium Incorporation into Other Substrates

We next evaluated the substrate promiscuity of SxtA AONS for deuterium incorporation. We anticipated that the other seven substrates for condensation (Gly, Ala, Leu, Ser, Trp, His, canavanine, see Chapter 3.3) would also be deuterated, as they must be deprotonated first in order

to form ketones. We screened twenty-four α-amino acids in total consisting of nineteen proteinogenic acids and five arginine-related acids (Figure 4.10A).

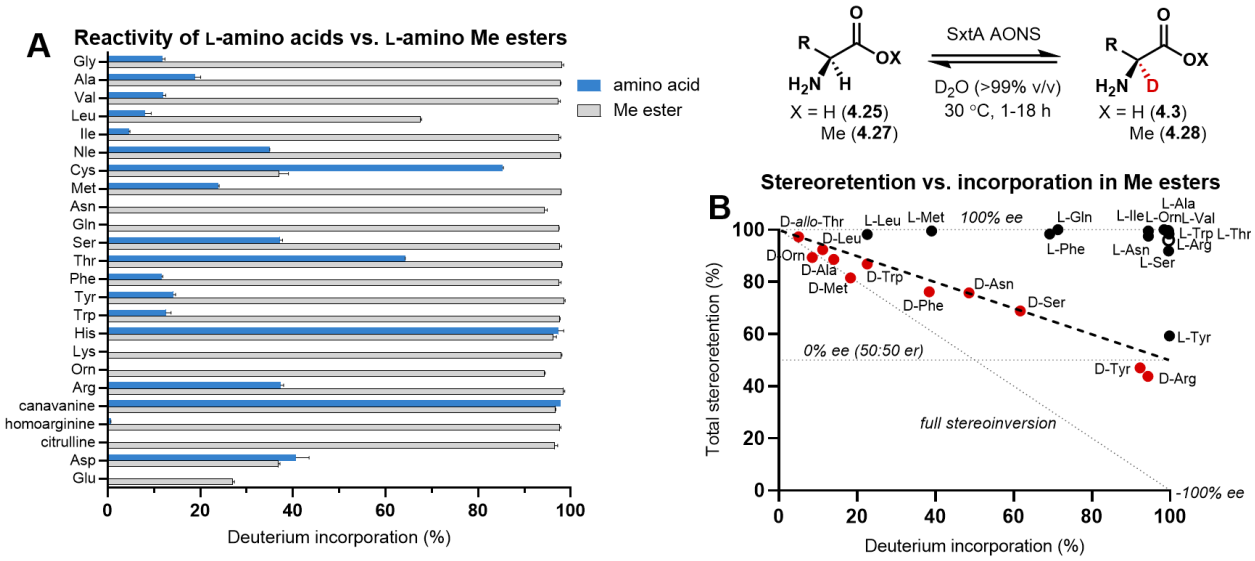


Figure 4.10. Deuterium incorporation at the alpha position of amino acids and their methyl esters.
(A) Deuterium incorporation into 4 mM of L-substrates after overnight incubation with SxtA AONS in D_2O. (B) Total stereoretention of 20 mM of the original methyl ester substrate configuration at the time of highest deuterium incorporation (see Section 4.6 Experimental for full data set).

In line with our initial studies with L-Arg (**L-4.19**) at 4 mM, most of these amino acids, including the structurally similar substrates L-ornithine, L-Lys and L-homoarginine, showed no or deuteration incorporation (Figure 4.10A, blue). Although some polar substrates demonstrated modest levels of deuterium incorporation, only two substrates were >90% deuterated, L-His and the Arg-mimic L-canavanine. The remaining five condensation substrates (Gly, Ala, Leu, Ser, Trp) were deuterated in low to modest levels. Based on our previous observation that esterification of Arg free acids to the corresponding methyl esters significantly increased both the rate and levels of deuterium incorporation (Figure 4.8C), we next assessed the corresponding methyl esters of the entire substrate panel (Figure 4.10A, gray). Most of these substrates were commercially available or accessible through synthesis in 1-2 steps. Gratifyingly, every L-methyl ester evaluated was

labeled to some extent with deuterium, and a majority (twenty side chains) were >90% labeled upon overnight incubation. The lowest-converting substrates were those with acidic side chains, as well as the polar ester L-Cys-OMe, and the aliphatic substrate L-Leu-OMe. Esterification slightly decreased the deuterium incorporation of Asp-based compounds, whereas L-Cys-OMe had significantly lower deuterium incorporation, likely due to the high levels of disulfide formation. Several D-methyl esters were also deuterated, but with no apparent correlation between the side chain and incorporation level (see Figure 4.S4). Thus, for SxtA AONS, a significantly greater breadth of methyl esters was deuterated compared to the analogous panel of amino acid substrates. The increased activity on methyl ester substrates may be due to the lower pK_a of the α-proton,[41,50,51] higher structural similarity to the ethyl ketone product **4.31**, or both.

To assess the stereocontrol in SxtA AONS-catalyzed deuteration of non-arginine esters, we resolved the stereoisomers of compounds possessing fifteen different side chains by chiral SFC (Figure 4.10B). In reactions with L-methyl esters (black dots), the L- configuration was retained, with many substrates affording both high deuterium incorporation and retention of the original configuration at the alpha position. L-Tyr-OMe was the only L-substrate observed to have relatively low stereoretention (41:59 D/L er). There is no significant correlation between the steric bulk or class of side chain and level of deuterium incorporation. For D-esters (red dots), deuterium incorporation varied widely from <1% in select aliphatic side chains (e.g., D-Val-OMe, see Table 4.S8 for full data set) to 94% for D-Arg-OMe. The mixtures of stereoisomers resulting from D-substrates are approximately halfway between full stereoretention and stereoinversion to L for every molecule processed. Thus, after deprotonation of a D-substrate and formation of a putative quinonoid intermediate, the reprotonation step appears to be non-selective (see Figure 4.8B). A racemic mixture of deuterated ester was also observed when SxtA AONS was replaced with free

PLP in nonenzymatic reactions; however, only low amounts of deuteration occur in the absence of enzyme (see Table 4.S2 for full data set).[52,53] For the achiral ester Gly-OMe, approximately 5% of the product mixture showed incorporation of two deuterium atoms. The remaining 95% of singly deuterated isotopomers could not be separated in our hands by chiral SFC analysis to determine enantioenrichment.

Condensation to Ethyl Ketones from Methyl Esters

The α-deuteration screen in Figure 4.10 demonstrated that seventeen of the amino acids that bind to the PLP cofactor of SxtA AONS and were deprotonated to form the first quinonoid intermediate (see **4.2**, Figure 4.6), the next step on-pathway to α-amino ketone formation. However, only ketones derived from eight of those seventeen acids (Arg, Gly, Ala, Leu, Ser, Trp, His, canavanine, see Chapter 3.3) were detected. Because all twenty-four of the methyl ester substrates screened bound and were deprotonated to generate deuterated products, the esters might also be converted to α-amino ketones or α-tetra-substituted compounds (Figure 4.11).

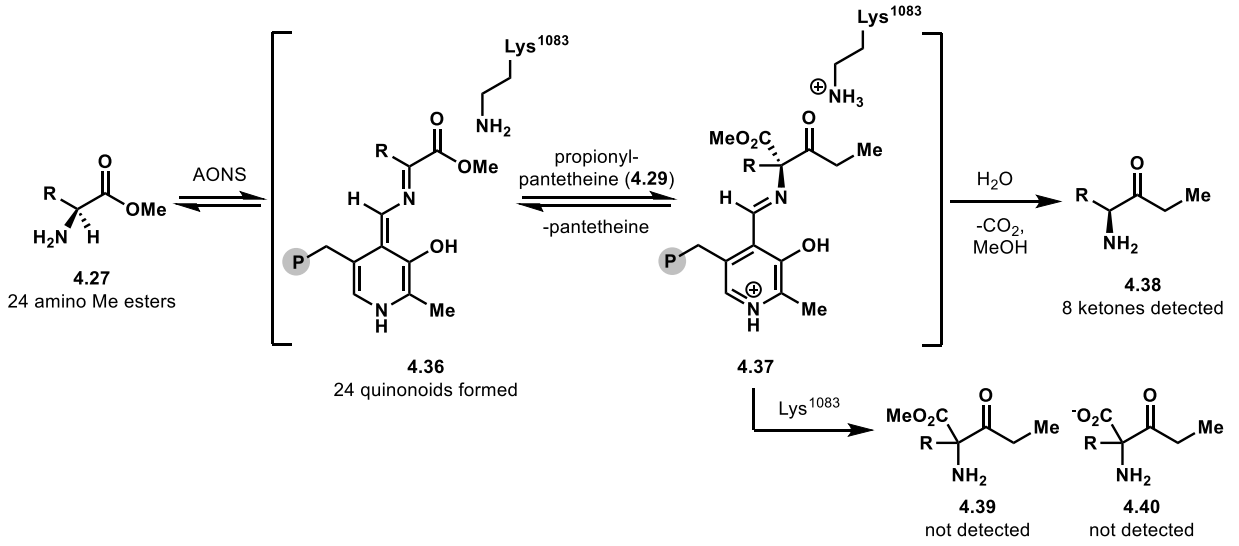


Figure 4.11. Potential ethyl ketone or α-tetra-substituted amino acid derivative formation from the corresponding methyl esters.

α-amino methyl esters and propionyl-pantetheine (**4.29**) were incubated with SxtA ACP-AONS and analyzed for formation of ethyl ketones (**4.38**) and α-tetra-substituted compounds by LC-MS. Only the same eight ketones were detected, but not quantified due to low overall conversion. We did not observe unhydrolyzed or undecarboxylated ketones (**4.39** and **4.40**). The esterification from α-amino acids to their corresponding methyl esters induced SxtA AONS to deprotonate all substrates, but the modification was not sufficient to continue to the addition of the propionyl group. Perhaps intermediate **4.36** more closely resembles quinonoid II, formed by deprotonation of PLP-ethyl ketone complexes (see **2.36** in Figure 2.7) rather than quinonoid I (see **4.2** in Figure 4.1). Formation of a diketone product from addition of another acyl group to the native product **4.31** has not previously been detected, and therefore, the protein conformation of quinonoid **4.36** may not be conducive to acyl addition. There may be another contributing factor in the ketone formation mechanism controlling the progression toward **4.38**. Spectroscopic studies of SxtA to identify some possible factors are detailed in Chapter 5.3.

Other AOS Enzymes

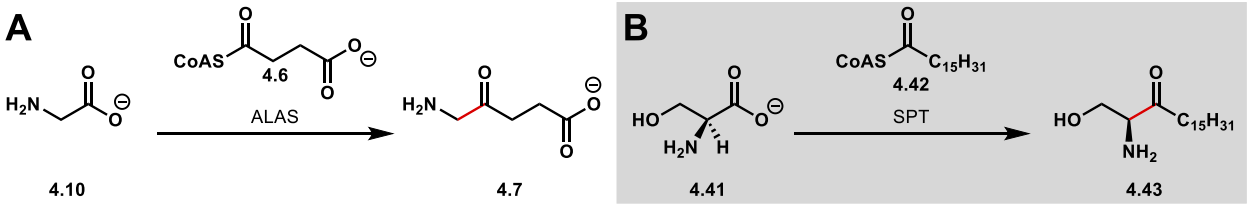


Figure 4.12. The native reactions of ALAS (A) and SPT (B).

We obtained two other AOS enzymes, *R. sphaeroides* ALAS and *Sphingomonas wittichii* serine palmitoyltransferase (SPT), which natively catalyze the reactions in Figure 4.12. The proteins were also screened for their ability to deuterate the α-position of α-amino acids. For ALAS (Figure 4.13, left), low deuterium incorporation was detected in a few amino acids (Met, Thr, Gln).

We observed modest incorporation into two substrates, Asn and Ser, and the remaining side chains had very minor activity. The native substrate Gly could not be detected on the Q-TOF MS instrument to confirm that *R. sphaeroides* ALAS also belongs to the group of AOS enzymes that deprotonates α-amino acids in the absence of its thioester substrate succinyl-CoA (**4.6**). Due to the acidic succinyl group added in the native condensation reaction (see **4.7**, Figure 4.12A), it was difficult to easily synthesize ester structural mimics to screen for increased deuteration activity as we did for SxtA AONS.

SPT's palmitoyl (C16) substrate is also aliphatic like SxtA AONS' propionyl group (see **4.43** and **4.44**, Figure 4.12B), so we screened the enzyme with both α-amino acids and their corresponding methyl esters (Figure 4.13, right). For the acids, side chains of similar size and polarity to the native Ser substrate (Cys, Met, Asn, Thr) were deuterated in modest to good amounts. Esterification generally increased incorporation, with significant improvement in deuterium levels in substrates like Gly-OMe and Ala-OMe. However, the gains in deuteration were not nearly as dramatic across the entire panel of SPT substrates as in the case for SxtA AONS. The lower pK_a of the α-amino ester compared to the acid may have contributed to the higher SPT activity, but we speculate that the rest of the acyl chain is also involved in shifting the protein's conformation to one favoring deprotonation.

We also intend to test α-amino myristyl esters (C14), closer structural mimics to the SPT product **4.44** as well. Even if this esterification does increase the deuteration, myristyl esters are more difficult to handle and prepare. The myristic alcohol used for esterification is a solid at room temperature and insoluble in water, whereas many of the methyl ester substrates in Figure 4.10 were synthesized neat in methanol. Therefore, we elected to use SxtA AONS over ALAS and SPT for preparative-scale reactions due to its higher activity and broader substrate scope.

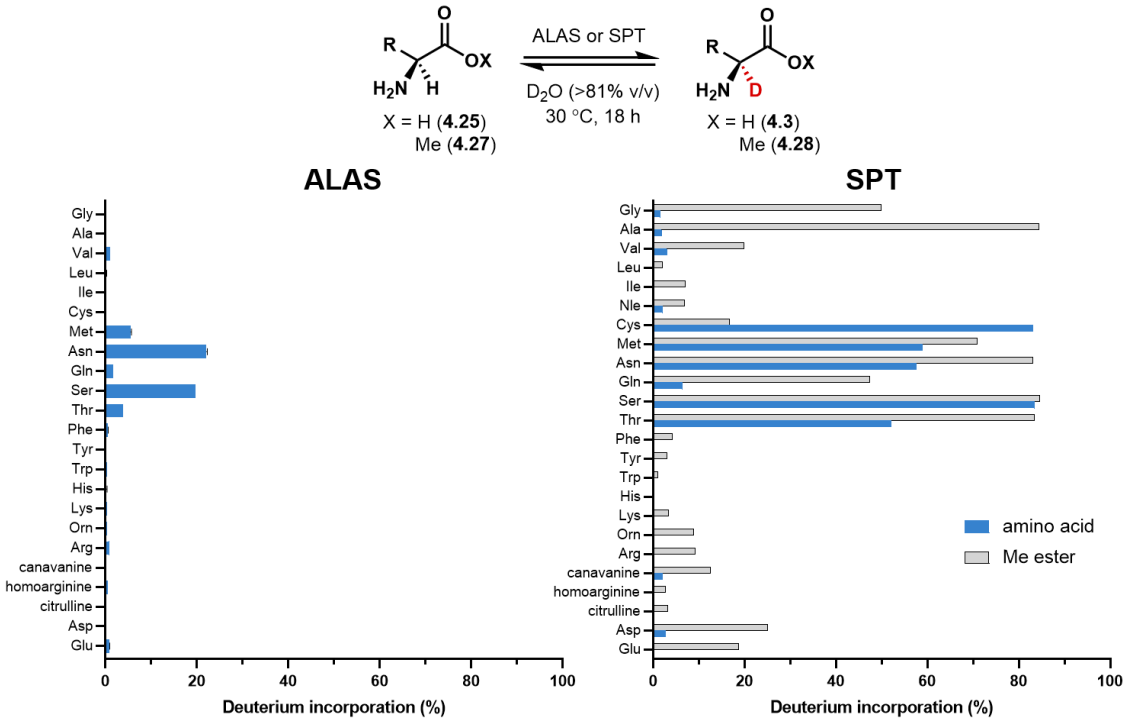


Figure 4.13. Deuterium incorporation into α-amino acids and methyl esters, catalyzed by AOS proteins ALAS and SPT.

4.4 Chemoenzymatic Synthesis of Deuterated Safinamide and Alanine

With evidence that SxtA AONS performs over a thousand turnovers, we sought to test the scalability of this transformation. As an initial target, we selected the Parkinson's drug, safinamide, which can be synthesized from an L-Ala-OMe precursor.[34,35,54]

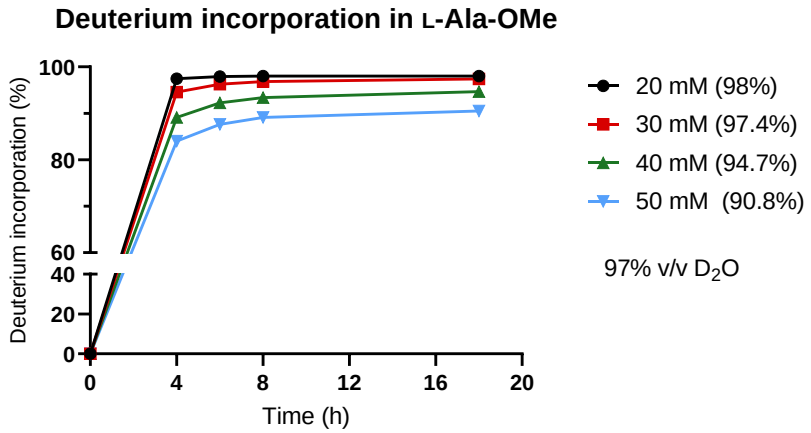


Figure 4.14. Deuterium incorporation into increasing concentrations of L-Ala-OMe.

We first maximized the substrate concentration, testing up to 50 mM ester and selected 40 mM for the best compromise between total deuterium incorporation and reaction volume (Figure 4.14). A 200-mg preparative scale deuteration reaction was conducted in 40 mL of D_2O, requiring nearly 40 mg of SxtA AONS, a quantity easily purified from 0.5 L of *E. coli* culture (Figure 4.15). Although high levels of hydrolysis of methyl esters to acids in phosphate buffer were previously observed, we conducted the scale-up in phosphate to minimize co-purification with HEPES. The HEPES buffer has a similar R_f by thin-layer chromatography to Ala-OMe (**4.44**) and is highly soluble in methanol, the organic solvent used to quench the enzymatic reaction and redissolve the ester for chromatography. We recovered 60% of the ester (**4.45**) from the deuteration reaction after purification by flash column chromatography with >99% deuterium incorporation and 96% ee. This deuterated building block was elaborated over three steps according to an established route to safinamide to deliver deutero-safinamide (**4.47**) with 96% final deuterium incorporation (Figure 4.15).[54]

To determine the final enantioenrichment, we also synthesized both protio-enantiomers of safinamide from D- and L-Ala-OMe. Although the primary amino group of Ala is now secondary in safinamide, both isomers proved to be too polar to be resolved by SFC. Literature on quality control of commercial-grade safinamide suggests that the drug compound, as well as the other very polar side chains we could not resolve by SFC in Figure 4.10, should be analyzed by ligand-exchange chromatography instead.[55,56] Assuming racemic reprotonation after deuterium loss throughout the safinamide synthesis, however, we estimate the final ee to be 93%.

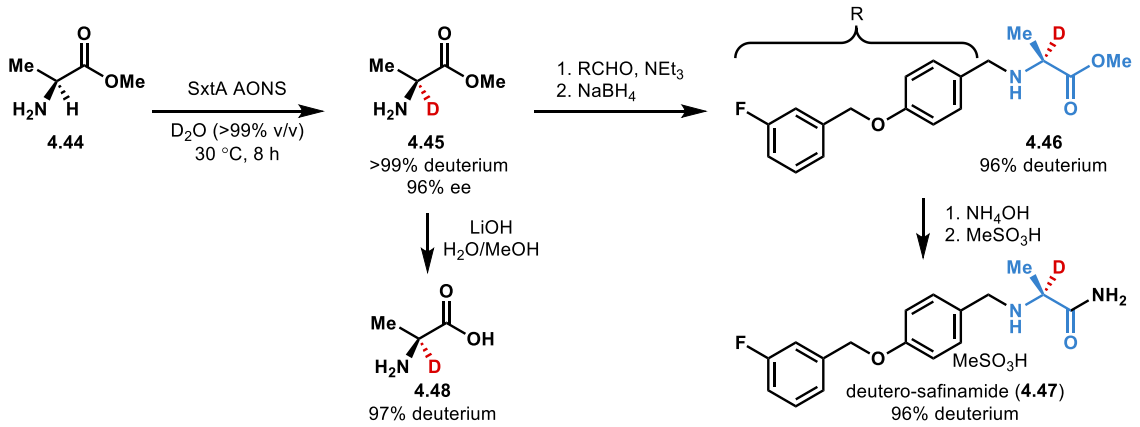


Figure 4.15. Synthesis of deutero-safinamide (4.47) and L-[2-^{2}H]alanine (4.48).

Deuterated ester **4.45** was also easily saponified with LiOH to α-deuterated alanine **4.48**, maintaining 97% of the α-deuterium label, an example of a building block useful in protein synthesis or other coupling reactions.[57,58]

4.5 Conclusions

We have demonstrated that SxtA AONS, a PLP-dependent enzyme that natively transforms an amino acid into an α-amino ketone, can be repurposed to install deuterium atoms site-selectively at the α-carbon of amino acids and amino esters. The presence of thioesters is not required for deuteration to occur. The method is particularly robust on a broad range of methyl ester substrates affording α-deuterated products under mild conditions. Even among promiscuous PLP-dependent amino acid-derivatizing enzymes, such as the broad-specificity racemases, SxtA AONS' ability to accommodate such a wide range of substrates that vary in structure and stereochemical configuration is unprecedented. The overall stereoselectivity of the C–D bond-forming step varies according to the side chain of α-amino methyl esters but in general, SxtA AONS exhibits high stereoselectivity with L-substrates, affording deuterated products with high levels of stereocontrol to generate products with the L- configuration. Finally, the utility and scalability of this one-step,

protecting group-free, biocatalytic method was applied to the chemoenzymatic synthesis of deuterium-labeled safinamide.

We anticipate that this method can easily be applied to the syntheses of other deuterated or tritiated products. Additionally, for substrates that are close to the line of full stereoinversion in Figure 4.10B (e.g., D-Ala-OMe, D-Met-OMe), SxtA AONS or its variants may potentially be leveraged for dynamic kinetic resolution. Furthermore, Ikushiro and coworkers have demonstrated that AOS enzymes in the "thioester required deprotonation" category can be used as biocatalysts for deuterium incorporation. Although deuterium incorporation occurred extremely slowly in L-Ser ($t_{1/2}$ = 103 ± 6 h) for *Sphingomonas paucimobilis* SPT, the rate was enhanced ($t_{1/2}$ = 0.75 ± 0.03 h) in the presence of nonhydrolyzable and unreactive thioether **4.49** (Figure 4.15), which mimics the structure of the native thioester substrate, palmitoyl-CoA (see **4.42**, Figure 4.12). Similarly, for SxtA AONS, binding of a pantetheine or CoA thioethers such as **4.51** could cause the desired protein conformational change to accelerate deuteration. New directions in biocatalytic α-deuteration mediated by wild-type SxtA AONS and the libraries currently being developed for α-amino ketone and α-tetra-substituted amino acid derivative synthesis will continue to be explored.

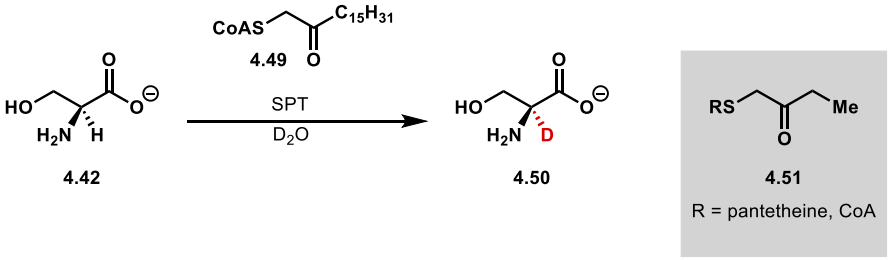


Figure 4.16. α-deuteration in a "thioester required" AOS enzyme by addition of a thioether that mimics the structure of the native thioester partner substrate.
See Ikushiro et al. [3]

The data from these deuterium incorporation experiments indicate that SxtA AONS can bind the majority of our L-amino acid substrate panel and deprotonate all twenty-four α-protons if the acid is esterified. However, only seven of the non-native side chains continue to be processed to α-amino ketones, so there must be multiple factors controlling possible conformational changes for deprotonation to the quinonoid intermediate and then addition to the thioester. The following Chapter will discuss our efforts to identify the residues and interactions responsible for SxtA AONS conformational changes.

4.6 Experimental

I. Chemical synthesis

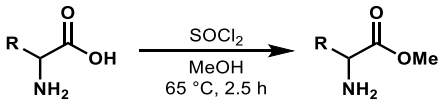
A. General information

All reagents were used as received unless otherwise noted. Reactions were carried out under a nitrogen atmosphere using standard Schlenck techniques unless otherwise noted. Solvents were degassed and dried over aluminum columns on an MBraun solvent system (Inert Corporation, Model PS-00-3). Reactions were monitored by thin layer chromatography using Machery-Nagel 60 F254 precoated silica TLC plates (0.25 mm) or Merck Silica Gel 60 RP-18 WF-254S precoated silica TLC plates (0.25 mm) which were visualized using UV, ninhydrin, *p*-anisaldehyde, cerium ammonium molybdate, 2,4-dinitrophenylhydrazine, or bromocresol green stain. Flash column chromatography was performed using Machery-Nagel 60 μm (230-400 mesh) silica gel. All compounds purified by flash column chromatography were sufficiently pure for use in further experiments unless otherwise indicated. Preparative-scale high performance liquid chromatography was carried out on a Waters 1525 Binary HPLC pump equipped with a Waters 2489 UV detector using a Phenomenex Kinetex 5 μm C18, 21.2 x 150 mm column. Solvent A = water with 0.1% v/v formic acid; solvent B = acetonitrile with 0.1% v/v formic acid. ^{1}H and ^{13}C NMR spectra were obtained in $CDCl_3$, CD_3OD or D_2O at rt (25 °C), unless otherwise noted, on Varian 400 MHz or Varian 600 MHz spectrometers. Chemical shifts of ^{1}H NMR spectra were recorded in parts per million (ppm) on the δ scale referenced to residual solvent peaks. Electrospray ionization liquid chromatography-mass spectrometry (LC-MS) analysis was performed on an Agilent G6545A quadrupole-time of flight mass spectrometer in positive mode with an Agilent 1290 UPLC system, or an Agilent 6320 time of flight mass spectrometer with an Agilent 1290 Infinity II UPLC system. Solvent A = water with 0.1% formic acid. Solvent B = 95% acetonitrile, 5% water and 0.1% formic acid. MS data were analyzed using the Agilent Qualitative Mass Hunter

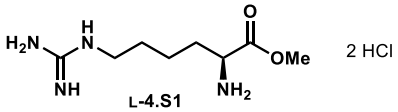
software. Chiral separation was performed on the Waters Investigator supercritical fluid chromatography (SFC) system. SFC data were analyzed using the Waters ChromScope Software.

B. Compound synthesis

General procedure for direct esterification of unprotected amino acids

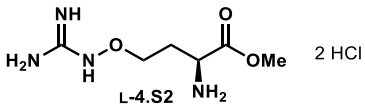


The amino acid (2.2 mmol, 1.0 equiv) was suspended in methanol (2.5 mL) in a 20 mL vial. Thionyl chloride (180 µL, 2.4 mmol, 1.1 equiv) was added dropwise to the suspension on ice, resulting in dissolution of the amino acid. The vial was then capped, and the reaction mixture stirred at 65 °C for 2.5 h. Diethyl ether (5 mL) was added. The reaction mixture was concentrated under reduced pressure and diluted with more ether (5 mL) repeatedly to remove residual acid until a foam or solid formed.

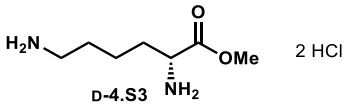


L-4.S1

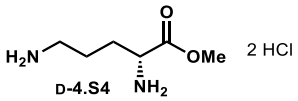
L-homoarginine methyl ester dihydrochloride (L-homoArg-OMe, L-4.S1): The general procedure afforded 610 mg of the title compound as a gray foam (quantitative yield). **^{1}H NMR** (600 MHz, D_2O) δ 4.15 (t, *J* = 6.4 Hz, 1H), 3.83 (s, 3H), 3.19 (t, *J* = 6.9 Hz, 2H), 2.05 – 1.87 (m, 2H), 1.67 – 1.59 (m, 2H), 1.55 – 1.35 (m, 2H). All spectra obtained were consistent with literature values.[59]



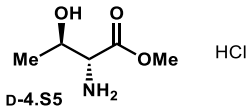
L-canavanine methyl ester dihydrochloride (L-4.S2): The general procedure afforded 96 mg of the title compound as an off-white foam (quantitative yield from 100 mg of L-canavanine sulfate). **^{1}H NMR** (600 MHz, D_2O) δ 4.31 (t, *J* = 6.6 Hz, 1H), 4.20 – 4.08 (m, 2H), 3.85 (s, 3H), 2.44 – 2.27 (m, 2H). All spectra obtained were consistent with literature values.[60]



D-lysine methyl ester dihydrochloride (D-4.S3): The general procedure afforded 520 mg of the title compound as a pale-yellow powder (quantitative yield). **^{1}H NMR** (600 MHz, D_2O) δ 4.15 (t, *J* = 6.5 Hz, 1H), 3.83 (s, 3H), 2.99 (t, *J* = 7.8 Hz, 2H), 2.08 – 1.87 (m, 2H), 1.70 (p, *J* = 7.8 Hz, 2H), 1.59 – 1.39 (m, 2H). All spectra obtained were consistent with literature values.[61]

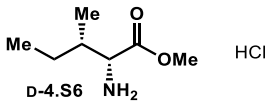


D-ornithine methyl ester dihydrochloride (D-4.S4): The general procedure afforded 480 mg of the title compound as a yellow powder (99% yield). **^{1}H NMR** (600 MHz, D_2O) δ 4.19 (t, *J* = 6.5 Hz, 1H), 3.85 (s, 3H), 3.04 (t, *J* = 7.7 Hz, 2H), 2.14 – 1.93 (m, 2H), 1.92 – 1.70 (m, 2H). All spectra obtained were consistent with literature values.[62]

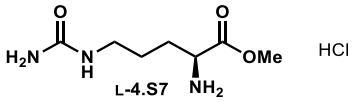


D-*allo*-threonine methyl ester hydrochloride (D-4.S5): The general procedure afforded 377 mg of the title compound as white needle crystals (quantitative yield). **^{1}H NMR** (600 MHz, CD_3OD)

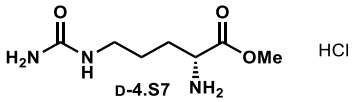
δ 4.25 (qd, *J* = 6.6, 3.6 Hz, 1H), 4.07 (d, *J* = 3.5 Hz, 1H), 3.85 (s, 3H), 1.26 (d, *J* = 6.6 Hz, 3H). All spectra obtained were consistent with literature values.[63]



D-*allo*-isoleucine methyl ester hydrochloride (D-4.S6): The general procedure afforded 400 mg of the title compound as a white foam (99% yield). **^{1}H NMR** (600 MHz, D_2O) δ 4.12 (d, *J* = 3.9 Hz, 1H), 3.83 (s, 3H), 2.18 – 2.09 (m, 1H), 1.54 – 1.41 (m, 1H), 1.37 – 1.22 (m, 1H), 0.96 (d, *J* = 7.1 Hz, 3H), 0.93 (t, *J* = 7.4 Hz, 3H). All spectra obtained were consistent with literature values.[64]

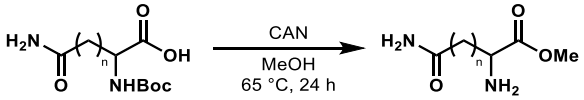


L-citrulline methyl ester hydrochloride (L-4.S7): The general procedure afforded 500 mg of the title compound as an off-white foam (quantitative yield). **^{1}H NMR** (600 MHz, D_2O) δ 4.15 (t, *J* = 6.4 Hz, 1H), 3.82 (s, 3H), 3.12 (td, *J* = 6.7, 3.0 Hz, 2H), 2.03 – 1.85 (m, 2H), 1.68 – 1.47 (m, 2H). All spectra obtained were consistent with literature values.[65]

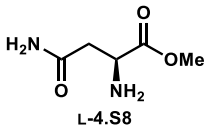


D-citrulline methyl ester hydrochloride (D-4.S7): The general procedure afforded 498 mg of the title compound as an off-white foam (99% yield). **^{1}H NMR** (600 MHz, D_2O) δ 4.15 (t, *J* = 6.4 Hz, 1H), 3.82 (s, 3H), 3.12 (td, *J* = 6.7, 3.0 Hz, 2H), 2.02 – 1.85 (m, 2H), 1.66 – 1.46 (m, 2H). All spectra obtained were consistent with literature values.[65]

General procedure for esterification and Boc-deprotection of substrates with amide-containing side chains (Asn, Gln). Adapted from Kuttan *et al*.[66]

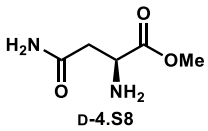


Boc-protected amino acid (2.6 mmol, 1.0 equiv) was resuspended in methanol (5.0 mL) in a 20 mL vial. Cerium ammonium nitrate (1.5 g, 2.6 mmol, 1.0 equiv) was added to the resuspension and the vial was sealed. The reaction stirred at 65 °C for 24 h, turning from a deep red suspension to a yellow solution overnight. Solvent was removed under reduced pressure and the crude residue was purified by flash chromatography (10% methanol v/v in dichloromethane).



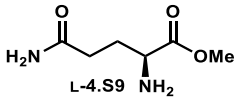
L-4.S8

L-asparagine methyl ester (L-4.S8): The general procedure afforded 411 mg of the title compound as off-white needle-shaped crystals (99% yield). R_f = 0.1 (10% v/v methanol in dichloromethane). **^{1}H NMR** (600 MHz, D_2O) δ 4.40 (dd, *J* = 6.1, 4.5 Hz, 1H), 3.83 (s, 3H), 3.08 (dd, *J* = 17.4, 6.1 Hz, 1H), 3.01 (dd, *J* = 17.4, 4.5 Hz, 1H). All spectra obtained were consistent with literature values.[67]

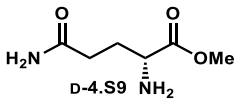


D-4.S8

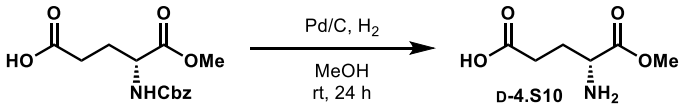
D-asparagine methyl ester (D-4.S8): The general procedure afforded 225 mg of the title compound as off-white needle-shaped crystals (54% yield). R_f = 0.1 (10% v/v methanol in dichloromethane). **^{1}H NMR** (600 MHz, D_2O) δ 4.41 (dd, *J* = 6.0, 4.5 Hz, 1H), 3.83 (s, 2H), 3.09 (dd, *J* = 17.4, 6.0 Hz, 1H), 3.01 (dd, *J* = 17.4, 4.5 Hz, 1H). All spectra obtained were consistent with literature values.[67]



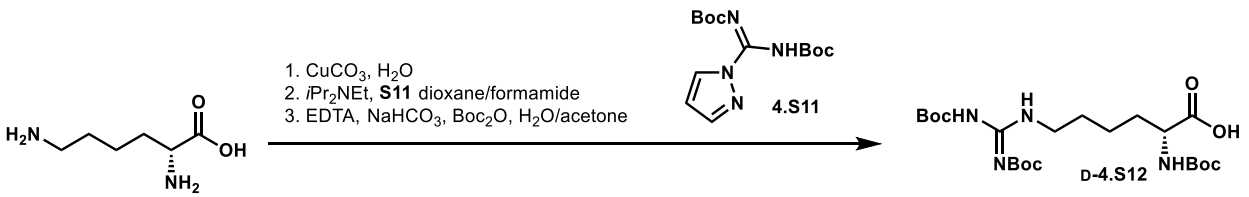
L-glutamine methyl ester (L-4.S9): The general procedure afforded 223 mg of the title compound as off-white needle-shaped crystals (50% yield). R_f = 0.1 (10% v/v methanol in dichloromethane). **^{1}H NMR** (600 MHz, D_2O) δ 4.17 (t, *J* = 6.7 Hz, 1H), 3.83 (s, 3H), 2.57 – 2.45 (m, 2H), 2.29 – 2.15 (m, 2H). All spectra obtained were consistent with literature values.[68]



D-glutamine methyl ester (D-4.S9): The general procedure afforded 353 mg of the title compound as off-white needle crystals (79% yield). R_f = 0.1 (10% v/v methanol in dichloromethane). **^{1}H NMR** (600 MHz, D_2O) δ 4.17 (t, *J* = 6.7 Hz, 1H), 3.83 (s, 3H), 2.57 – 2.44 (m, 2H), 2.29 – 2.14 (m, 2H). All spectra obtained were consistent with literature values.[68]

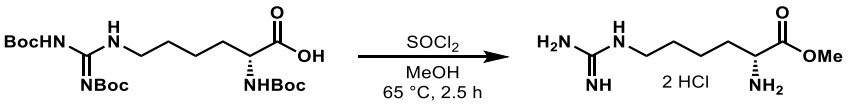


D-glutamic acid α-methyl ester (D-4.S10): A round-bottomed flask was charged *N*-Cbz-D-glutamic acid alpha-methyl ester (200 mg, 0.68 mmol), 10 wt% Pd/C (50 mg) and methanol (10 mL). The flask was sparged with N_2 for 20 min then equipped with an H_2-filled balloon and sparged again for 15 min. The reaction was stirred under an H_2 atmosphere for 24 h. The reaction mixture was filtered through celite and concentrated to afford 106 mg of the title compound as a white powder (97% yield). **^{1}H NMR** (600 MHz, D_2O) δ 4.14 (t, *J* = 6.4 Hz, 1H), 3.82 (s, 3H), 2.37 (td, *J* = 7.3, 2.6 Hz, 2H), 2.23 – 2.06 (m, 2H). All spectra obtained were consistent with the commercially available L-glutamic acid α-methyl ester.[69]



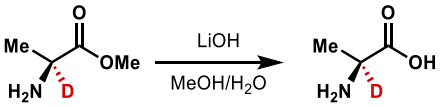
Tri-Boc-protected D-homoarginine (D-4.S12): procedure adapted from Tsuchiya et al.[70] D-lysine hydrochloride (66 mg, 0.36 mmol, 1.0 equiv) and anhydrous copper carbonate (160 mg, 0.72 mmol, 2.0 equiv) were resuspended in 2.5 mL of water in a 20 mL vial equipped with a stir bar. The vial was sealed and the reaction mixture heated at 100 °C for 1 h. The cooled reaction mixture was then filtered to remove the insoluble copper salt. The blue filtrate was concentrated to a residue in a 10 mL round bottom flask equipped with a stir bar and resuspended in 350 µL of dioxane and 717 µL of formamide. **4.S11** (110 mg, 0.36 mmol, 1.0 equiv) and *i*Pr$_2$NEt (150 µL, 0.83 mmol, 2.3 equiv) were added to the resuspension, which stirred overnight under N_2 at rt. The following day, the clear blue solution was diluted with EtOAc (5 mL) and brine (10 mL), and extracted with EtOAc (3 x 15 mL). The homoarginine intermediate was not especially soluble in EtOAc, so the combined organic layers were not dried, but concentrated slowly down to a pale blue powder in a 25 mL round bottom flask. To the flask was added acetone (5 mL), water (1 mL), EDTA·2Na (88 mg, 0.24 mmol, 0.66 equiv), $NaHCO_3$ (93 mg, 1.1 mmol, 3.0 equiv) and Boc anhydride (180 mg, 0.83 mmol, 2.3 equiv), resulting in an opaque blue viscous resuspension that stirred vigorously overnight under N_2 at rt. After 16 h, the reaction mixture had become biphasic—a clear organic layer on top and a clear blue aqueous lower layer. The mixture was diluted with brine (5 mL) and extracted with EtOAc (3 x 15 mL). The combined organic layers were dried over Na_2SO_4, filtered and concentrated to a pale green residue. Flash chromatography (5% v/v methanol in dichloromethane) afforded 106 mg (60% yield over three steps) of the title compound as a clear residue. R_f = 0.3. **^{1}H NMR** (600 MHz, CD_3OD) δ 4.07 (dd, *J* = 9.2, 4.6 Hz, 1H), 3.37 (dt, *J* = 7.1, 2.7 Hz, 2H), 1.85 (m, 2H), 1.72 – 1.56 (m, 4H), 1.54 – 1.51 (m, 9H), 1.47 (m, 9H), 1.45 (m, 9H);

^{13}C NMR (150 MHz, CD_3OD) δ 164.4, 164.3, 158.1, 157.5, 154.1, 84.4, 80.4 (2 carbons), 49.9, 41.6, 32.6, 29.8, 28.8, 28.6, 28.3, 24.2; **HRMS** (ESI) m/z calcd for $C_{22}H_{41}N_4O_8^+$ $[M+H]^+$ 489.2919 m/z, found 489.2922. All spectra obtained were consistent with literature values.[71]



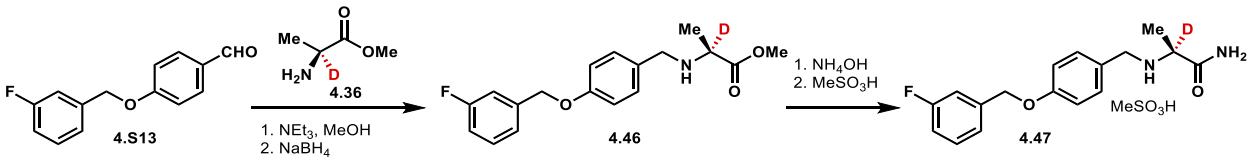
D-4.S12 D-4.S1

D-homoarginine methyl ester dihydrochloride (D-4.S1): The general procedure for direct esterification of unprotected amino acids afforded 35 mg (quantitative yield) of the title compound from tri-Boc-protected D-homoarginine **D-S12** (43 mg) obtained as a reddish-brown residue. **^{1}H NMR** (600 MHz, D_2O) δ 4.14 (t, *J* = 6.4 Hz, 1H), 3.83 (s, 3H), 3.18 (t, *J* = 6.9 Hz, 2H), 2.03 – 1.86 (m, 2H), 1.65 – 1.58 (m, 2H), 1.55 – 1.28 (m, 2H). All spectra obtained were consistent with literature values.[59]

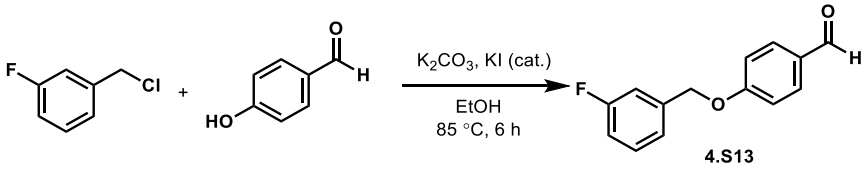


4.45 4.48

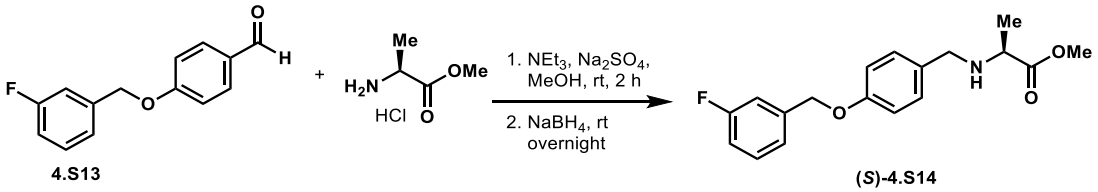
L-[2-^{2}H]alanine (4.48): L-[2-^{2}H]alanine methyl ester **4.35** (20 mg, 0.19 mmol, 1.0 equiv) as a 1 M solution in methanol, and LiOH (9 mg, 0.38 mmol, 2.0 equiv) as a 1 M solution in water were combined and stirred in a 4 mL vial at rt for 30 min (monitored by TLC, 10% v/v methanol in dichloromethane, starting material R_f = 0.2). The solvent was then removed under reduced pressure, affording a white powder (21 mg as a mixture of lithium salts). R_f = 0. **^{1}H NMR** (400 MHz, D_2O) δ 1.20 (s, 3H); **^{13}C NMR** (150 MHz, D_2O) δ 185.3, 51.7 (t, *J* = 21 Hz), 20.9; **HRMS** (ESI) m/z calcd for $C_3H_7DNO_2$ $[M+H]^+$ 91.0612 m/z, found 91.0616 (97.6% deuterium incorporation).[72]



Deuterated safinamide and the (*R*)-enantiomer were synthesized according to a procedure modified from Du et al.[54]

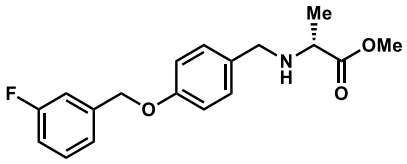


4-((3-fluorobenzyl)oxy)benzaldehyde (4.S13): To a flame-dried 25 mL flask equipped with a stir bar were added 1-(chloromethyl)-3-fluorobenzene (1.2 mL, 9.6 mmol, 1.2 equiv), 4-hydroxybenzaldehyde (0.98 g, 8.0 mmol, 1.0 equiv), anhydrous potassium carbonate (1.1 g, 8.0 mmol, 1.0 equiv), potassium iodide (0.13 mg, 0.8 mmol, 0.1 equiv) and anhydrous ethanol (12 mL). The reaction mixture refluxed at 85 °C for 6 h, and then partially concentrated to a slurry. The slurry was diluted with brine and extracted with EtOAc. The combined organic layers were washed with sat. aq. sodium thiosulfate to remove excess iodide, then dried over Na_2SO_4, filtered and concentrated. Flash chromatography on silica gel (25% v/v EtOAc in hexanes) afforded 1.68 g (91% yield) of the title compound as an off-white solid. R_f = 0.3. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 9.90 (s, 1H), 7.89 – 7.81 (m, 2H), 7.37 (td, *J* = 7.9, 5.8 Hz, 1H), 7.22 – 7.13 (m, 2H), 7.06 (dd, *J* = 8.2, 6.4 Hz, 3H), 5.15 (s, 2H). All spectra obtained were consistent with literature values.[73]



Methyl (4-((3-fluorobenzyl)oxy)benzyl)-L-alaninate ((*S*)- 4.S14): To a flame-dried 10 mL flask equipped with a stir bar were added aldehyde **4.S13** (166 mg, 0.72 mmol, 1.0 equiv), L-alanine

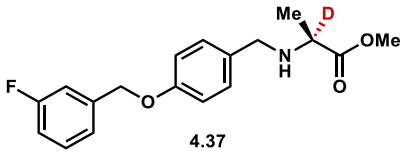
methyl ester hydrochloride (100 mg, 0.72 mmol, 1.0 equiv), Na_2SO_4 (36 mg, 0.25 mmol, 0.35 equiv), and anhydrous methanol (0.9 mL). NEt_3 was added dropwise and the reaction stirred under N_2 at rt for 2 h. The reaction mixture was then filtered to remove solids. $NaBH_4$ (28 mg, 0.72 mmol, 1.0 equiv) was added in two portions to the filtrate, which bubbled and then formed a white suspension overnight. The following day, the reaction mixture was concentrated slightly under reduced pressure, diluted with brine (5 mL) and extracted with toluene (3 x 15 mL). The combined organic layers were dried over Na_2SO_4, filtered and concentrated. The residue was purified by flash chromatography on silica gel (50% v/v EtOAc in hexanes), affording 200 mg (86% yield) of the title compound as a clear oil. R_f = 0.3. **^{1}H NMR** (600 MHz, CD_3OD) δ 7.37 (td, *J* = 8.0, 5.8 Hz, 1H), 7.28 – 7.20 (m, 3H), 7.17 (d, *J* = 9.5, 1H), 7.06 – 6.99 (m, 1H), 6.98 – 6.90 (m, 2H), 5.09 (s, 2H), 3.70 (s, 3H), 3.68 (d, *J* = 12.6 Hz, 1H), 3.60 (d, *J* = 12.6 Hz, 1H), 3.36 (q, *J* = 7.0 Hz, 1H), 1.29 (d, *J* = 7.0 Hz, 3H). All spectra obtained were consistent with literature values.[54]



(*R*)-4.S14

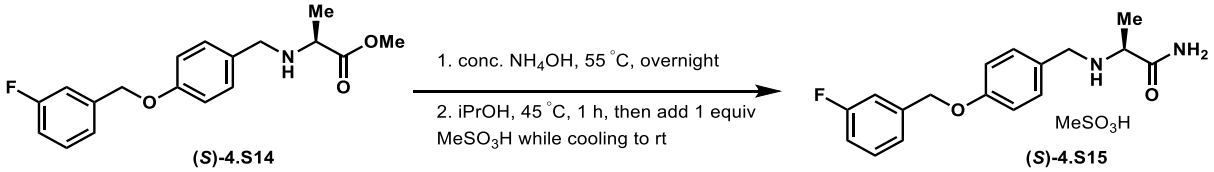
Methyl (4-((3-fluorobenzyl)oxy)benzyl)-D-alaninate ((*R*)- 4.S14): The above procedure afforded 123 mg (54% yield) of the title compound as a clear oil from 100 mg of D-alanine methyl ester hydrochloride.**^{1}H NMR** (600 MHz, CD_3OD) δ 7.33 (td, *J* = 7.9, 5.8 Hz, 1H), 7.26 – 7.22 (m, 2H), 7.20 (d, *J* = 7.6 Hz, 1H), 7.14 (d, *J* = 9.9 Hz, 1H), 7.00 (td, *J* = 8.6, 2.7 Hz, 1H), 6.94 – 6.90 (m, 2H), 5.03 (s, 2H), 3.67 (overlapping singlet and doublet, *J* = 14.7 Hz, 4H), 3.59 (d, *J* = 12.6 Hz, 1H), 3.37 (q, *J* = 7.0 Hz, 1H), 1.28 (d, *J* = 7.0 Hz, 3H); **^{13}C NMR** (150 MHz, CD_3OD) δ 176.5, 164.3 (d, *J* = 245 Hz), 159.2, 141.6 (d, *J* = 7 Hz), 132.4, 131.3 (d, *J* = 8 Hz), 131.0, 124.0 (d, *J* =

3 Hz), 115.9, 115.4 (d, *J* = 21 Hz), 115.0 (d, *J* = 22 Hz), 69.9, 56.4, 52.4, 51.8, 18.5; **HRMS** (ESI) m/z calcd for $C_{18}H_{21}FNO_2$ $[M+H]^+$ 318.1500 m/z, found 318.1503.



4.37

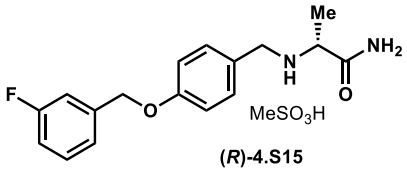
Methyl (4-((3-fluorobenzyl)oxy)benzyl)-L-alaninate-2-*d* (4.46): The above procedure afforded 34 mg (58% yield) of the title compound as a clear oil from 20 mg (0.19 mmol) of **4.45.** **^{1}H NMR** (600 MHz, CD_3OD) δ 7.36 (td, *J* = 7.9, 5.8 Hz, 1H), 7.29 – 7.20 (m, 3H), 7.16 (dt, *J* = 9.9, 2.1 Hz, 1H), 7.02 (td, *J* = 8.6, 2.6 Hz, 1H), 6.97 – 6.91 (m, 2H), 5.07 (s, 2H), 3.69 (s, 3H), 3.67 (d, *J* = 12.6 Hz, 1H), 3.58 (d, *J* = 12.6 Hz, 1H), 1.27 (s, 3H); **^{13}C NMR** (150 MHz, CD_3OD) δ 176.8, 164.4 (d, *J* = 245 Hz), 159.3, 141.7 (d, *J* = 7 Hz), 132.7, 131.3 (d, *J* = 8 Hz), 131.0, 124.0 (d, *J* = 3 Hz), 115.9, 115.4 (d, *J* = 21 Hz), 115.0 (d, *J* = 22 Hz), 70.0, 56.1 (t, *J* = 21 Hz), 52.4, 51.9, 18.5; **HRMS** (ESI) m/z calcd for $C_{18}H_{20}DFNO_2$ $[M+H]^+$ 319.1563 m/z, found 319.1568 (96% deuterium incorporation).



(*S*)-4.S14 (*S*)-4.S15

(*S*)-safinamide, methanesulfonate salt ((*S*)- 4.S15): To a 20 mL vial equipped with a stir bar was added ester **(*S*)-4.S14** (150 mg, 0.46 mmol, 1.0 equiv) and 2 mL of concentrated aqueous ammonia. The vial was sealed and heated at 55 °C overnight. After 16 h, a white resuspension had formed. The reaction mixture was concentrated to a white powder (140 mg) under reduced pressure, 50 mg of which was purified by preparative high-performance liquid chromatography. Solvent A = water with 0.1% formic acid. Solvent B = acetonitrile with 0.1% formic acid; method = 5% B at 10

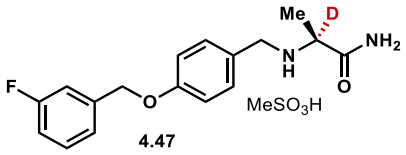
mL/min for 1 min, followed by a linear gradient to 70% B over 7 min, and a second linear gradient to 100% B over 7 min, 100% B for 1 min, a linear gradient to 5% B over 1 min and then 1 min re-equilibration at 5% B (total time 18 min). The desired amide product eluted at 7.7 min, partially separated from a major side product (RT = 8.5 min) that appears to be the carboxylic acid. Relevant fractions were combined and concentrated to a white powder (35 mg). To the powder was added 5 mL of isopropanol, and the mixture was heated to 45 °C. Methanesulfonic acid (7.6 µL, 0.12 mmol, 1.0 equiv) was added to the reaction mixture, which was cooled slowly to rt to precipitate the product salt. The suspension was concentrated, affording 41 mg (approx. 70% of the amount chromatographed) of the title compound as a white powder. **^{1}H NMR** (600 MHz, CD_3OD) δ 7.46 – 7.41 (m, 2H), 7.38 (td, *J* = 8.0, 5.8 Hz, 1H), 7.24 (dt, *J* = 7.6, 1.3 Hz, 1H), 7.17 (dt, *J* = 9.8, 2.1 Hz, 1H), 7.10 – 7.05 (m, 2H), 7.03 (td, *J* = 8.5, 2.4 Hz, 1H), 5.14 (s, 2H), 4.11 (s, 2H), 3.95 (q, *J* = 7.0 Hz, 1H), 2.71 (s, 3H), 1.55 (d, *J* = 7.0 Hz, 3H). All spectra obtained were consistent with literature values.[55]



(*R*)-4.S15

(*R*)-safinamide, methanesulfonate salt ((*R*)- 4.S15): The above procedure afforded 17 mg of the title compound as a white powder from starting from 50 mg of **(*R*)-4.S14** (27% over 2 steps). **^{1}H NMR** (600 MHz, CD_3OD) δ 7.45 – 7.41 (m, 2H), 7.38 (td, *J* = 8.0, 5.8 Hz, 1H), 7.25 (dt, *J* = 7.7, 1.3 Hz, 1H), 7.17 (dt, *J* = 9.8, 2.1 Hz, 1H), 7.10 – 7.06 (m, 2H), 7.03 (td, *J* = 8.4, 2.3 Hz, 1H), 5.15 (s, 2H), 4.11 (s, 2H), 3.93 (q, *J* = 7.0 Hz, 1H), 2.71 (s, 3H), 1.55 (d, *J* = 7.0 Hz, 3H); **^{13}C NMR** (150 MHz, CD_3OD) δ 172.4, 164.4 (d, *J* = 245 Hz), 160.9, 141.3 (d, *J* = 7 Hz), 132.8, 131.4 (d, *J* = 8 Hz), 124.6, 124.1 (d, *J* = 3 Hz), 116.6, 115.6 (d, *J* = 21 Hz), 115.0 (d, *J* = 22 Hz), 70.0, 56.4,

50.5, 39.5, 16.9; **HRMS** (ESI) m/z calcd for $C_{17}H_{20}FN_2O_2$ $[M+H]^+$ 303.1503 m/z, found 303.1512. All spectra obtained were consistent with literature values.[55]



4.47

Deuterated safinamide, methanesulfonate salt (4.47): The above procedure afforded 13 mg (32% yield over 2 steps) of the title compound as a white powder from 32 mg (0.10 mmol) of **4.46**. **^{1}H NMR** (600 MHz, CD_3OD) δ 7.45 – 7.41 (m, 2H), 7.38 (td, *J* = 8.0, 5.8 Hz, 1H), 7.26 – 7.23 (m, 1H), 7.17 (dt, *J* = 9.8, 2.1 Hz, 1H), 7.10 – 7.06 (m, 2H), 7.06 – 7.01 (m, 1H), 5.15 (s, 2H), 4.10 (s, 2H), 2.71 (s, 3H), 1.54 (s, 3H); **^{13}C NMR** (150 MHz, CD_3OD) δ 172.3, 164.4 (d, *J* = 245 Hz), 160.9, 141.3 (d, *J* = 7 Hz), 132.8, 131.4 (d, *J* = 8 Hz), 124.6, 124.1 (d, *J* = 3 Hz), 116.6, 115.6 (d, *J* = 21 Hz), 115.0 (d, *J* = 22 Hz), 70.0, 56.1 (t, *J* = 21 Hz), 50.4, 39.5, 16.8; **HRMS** (ESI) m/z calcd for $C_{17}H_{19}DFN_2O_2$ $[M+H]^+$ 304.1566 m/z, found 304.1565 (95.8% deuterium incorporation). Due to the polarity of safinamide, the final ee could not be determined by SFC.

II. Cloning, protein expression and purification

A. General information

Escherichia coli cloning strain DH5α (Invitrogen) was used for DNA propagation. Protein was expressed in *E. coli* strains BL21(DE3) (New England BioLabs) after transformation with *MW_sxtA_AONS* subcloned into pMCSG7K (provided by the University of Michigan Center for Structural Biology) plasmids. Cells were grown in Luria Bertani (LB) or Terrific broth (TB) with 4% v/v glycerol supplemented with kanamycin (50 µg/mL) purchased from Gold Biotechnology.

A pET28b plasmid containing *Swit_SPT* was purchased from Twist Biosciences.

Primers were purchased from Integrated DNA Technologies. DpnI restriction enzyme was purchased from New England BioLabs. Ligation-independent cloning (LIC)-qualified T4 DNA polymerase was purchased from EMD Millipore. QIAquick PCR purification, gel extraction and miniprep kits were purchased from Qiagen. HisPur nickel-nitrilotriacetic acid (Ni-NTA) resin was purchased from Thermo Scientific. Proteins were concentrated using Amicon centrifugal filters purchased from EMD Millipore at 4,000 x *g*, 4 °C. PD-10 desalting columns were purchased from GE Healthcare. Protein samples were analyzed on Mini-PROTEAN TGX Gels (4-15%) from BioRad and visualized with Protein Ark Quick Coomassie Stain from Anatrace. Proteins were quantified with the Pierce 660 nm Assay Reagent from Thermo Scientific.

B. Sequence information

pMCSG7K *Microseira wollei sxtA* AONS *E. coli*-optimized sequence. GenBank accession ACG63826.1; linkers and the 6x-His tag are shown in bold.

atgcaccatcatcatcattcttctggtgtagatctgggtaccgagaacctgtacttccaatccaatgcaGAAAA
GCAGAGTGATCTGGCGAGCCCCGTGGACTTAGCTAATGAAATCAACACATTAGAAAATGTTGTCAATCAGCAGAAGA
TCCCTCAGGTTACGCGCGTAGTGACTGAACAGCAAGGTCGTAAGGTCTTGATCGATGGCCACTGGGTGATTGACTTC
GCCTCATGCAACTATTTAGGTTTAGACTTGCACCCCAAAGTGAAGGAGGCTATCCCCCCGGCTTTGGAGAAGTGGGG
GGTCCATCCCTCTTGGACACGTTTGGTGGCCAGCCCGGTAATCTATGAGGAATTAGAGGAGGAGTTAAGCAAACTGT
TAGGAGTTCCCGACGTATTAGTTTTCCCAAGCGTCACATTACTTCAGATCGGGGTCTTGCCACTTCTGACAGGCTCG
AACGGAGTTATTTTAGGCGACATTGCTGCGCATCGTTGCATTTACGAGGCCTGCTGCCTGGCACAGCAGAAGGGCTC
ACAGTTTATCCAATATCGCCACAATGACTTGAATGATTTAGCGGAGAAATTGGCAAAGTATCCACTGGAACAAGTCA
AAATCATCGTCATTGACGGAGTGTACTCTATGTCCGCCGACTTCCCAGACCTTCCGGCTTACGTCCGCCTGGCGAAG
GAATACAATGCTTTTATTTATATGGATGACGCGCACGGATTTGGTATTTTGGGTGAGAACCCGTCCTCCGATATGCC
TTACGGATACAAAGGTAATGGTATGGTGAATTATTTCGATCTTCAATTTGCCGAGGACAACATTATCTACGTCGCGG
GGTTGTCTAAAGCCTACAGTTCGTACGCCGCCTTTCTTACTTGCGGTGATCGTCAGATTAAGACCAATTTCCGTAAT
GCTTGGACTGCCATTTTCAGCGGTCCTAGCCCAGTGGCAAGCTTGGCATCAGCCTTGGCCGGCTTACAGGTTAATCG
TCAGGAAGGAGAACAACTTCGCAAGCAAATCTACCACTTGACGCGCAAATTGGTGACCCAAGCTCGCGCCATTGGCT
TCGAAGTGGACAACTACGGTTACGTTCCCATCGTTTCCGTCTTAGTCGGAGACGCGCAACACATGATTGACGTATGT
CAGCTGCTGTGGGAGTACGGCATTCTGATCACACCCGCTATTTTCCCAATTGTTCCACTTAATAAATCCGCGCTGCG
CTTTTCTATCACAGCTGCCAATACGGAAGAAGAGATTGACCAAGCCATCAAAGCATTGGAGGGGGTATGGGTCCTGC
TTCAGAAAGAGAAG**cattag**

Translation: linkers and the 6x-His tag are shown in bold.

MHHHHHHSSGVDLGTENLYFQSNAEKKSDLASPVDSANEINTLENVVNQQKIPQVTRVVTEQQDRKVLIDGHWVIDF
ASCNYLGLDLHPKVKEAIPPALEKWGVHPSWTRLVASPAIYEELEEELAKLLGVPDVLLFPSLTLLQMGVLPLLTGN
NGVIFGDISAHRCIYEACCLAQHKGAQFIQYRHNDLNDLAAKLAKYPLEQVKIIAIDGVYSMSADLPDLPAYVRLAK
QYNALIYIDDAHGFGILGENPSSDMPYGYKGNGIVNYFDLRFAEDNIIYVAGLSKAYSSYAAFVTCGNSQIKTQFRN
AWTAIFSGPSPVASLASALAGLQVNRQEGEQLRKQVYYLTHKLVTQARAIGFEVDNYSYVPIVSVLVGDAQHIVDVC
QLLWEYGILITPAIFPIVPLNKTALRFSITAANTEEEIDQAIKALEAVWDLLQKRKALLCK**H***

Primers used: overhangs for ligation-independent cloning are shown in boldface. The reverse primer encodes an extra His residue at the C terminus as an error in the primer design (underlined).

MW_AONS_LIC_fwd – 5′-**TACTTCCAATCCAATGCA**GAAAAGAAGTCCGATCTGGC-3′

MW_AONS_LIC_rev – 5′-**TTATCCACTTCCAATGCTA**ATGCTTACATAAAAGGGCTTTGCG-3′

pMCSG7 *Rhodobacter capsulatus hemA* (aminolevulinic acid synthase, ALAS) *E. coli*-optimized sequence. GenBank accession ADE85192.1; linkers and the 6x-His tag are shown in bold.

atgcaccatcatcatcattcttctggtgtagatctgggtaccgagaacctgtacttccaatccaatgcaatgGA
TTACAACTTGGCATTAGATAAAGCCATCCAAAAGTTGCATGACGAAGGCCGCTATCGTACCTTCATCGACATTGAGC
GTGAAAAAGGAGCTTTTCCTAAAGCGCAGTGGAATCGTCCCGATGGAGGGAAACAAGACATCACTGTCTGGTGCGGC
AACGATTATTTGGGCATGGGTCAACACCCTGTGGTCCTTGCCGCCATGCACGAGGCCTTAGAGGCCGTCGGAGCTGG
GTCTGGAGGCACACGCAACATTAGTGGCACGACCGCATATCATCGTCGTCTTGAGGCAGAAATTGCCGATTTGCACG
GGAAGGAGGCAGCTTTGGTCTTCTCGTCTGCTTACATTGCGAATGATGCGACGCTGAGCACGTTACGTGTCCTTTTT
CCAGGGCTTATCATTTATAGTGACTCATTGAATCATGCGTCTATGATCGAAGGGATCAAGCGTAATGCAGGGCCGAA
ACGCATCTTCCGTCACAACGACGTTGCTCACCTTCGTGAATTGATTGCGGCAGATGACCCGGCGGCCCCAAAATTAA
TTGCGTTTGAGTCCGTTTACAGTATGGATGGCGATTTCGGACCTATTAAGGAGATCTGCGATATTGCGGACGAGTTT
GGTGCCCTTACATACATCGATGAAGTACATGCAGTAGGGATGTACGGCCCTCGCGGTGCAGGAGTCGCGGAGCGCGA
TGGGCTGATGCACCGTATCGACATCTTCAATGGAACCTTGGCGAAAGCATACGGCGTATTCGGCGGCTATATTGCTG
CTTCGGCGAAAATGGTGGATGCCGTTCGTTCCTATGCGCCGGGGTTTATTTTCTCAACTTCGCTTCCGCCTGCGATC
GCTGCTGGAGCCCAAGCGTCTATCGCGTTTTTGAAAACCGCCGAAGGACAGAAATTACGCGACGCGCAACAAATGCA
CGCAAAGGTCCTGAAAATGCGCCTTAAAGCCCTTGGGATGCCCATCATTGATCACGGATCCCATATTGTGCCTGTCG
TAATCGGGGATCCCGTGCACACTAAAGCCGTGAGCGATATGTTGCTTTCGGATTATGGGGTTTATGTCCAACCAATC
AACTTTCCCACGGTCCCACGCGGAACGGAGCGTCTTCGCTTTACCCCGAGCCCTGTCCATGACTTAAAGCAGATTGA
CGGGTTGGTGCACGCCATGGATCTTCTTTGGGCTCGTTGTGCACTTAATCGTGCAGAAGCCTCGGCT

Translation: linkers and the 6x-His tag are shown in bold.

MHHHHHHSSGVDLGTENLYFQSNAMDYNLALDKAIQKLHDEGRYRTFIDIEREKGAFPKAQWNRPDGGKQDITVWCG
NDYLGMGQHPVVLAAMHEALEAVGAGSGGTRNISGTTAYHRRLEAEIADLHGEAALVFSSAYIANDATLSTLRVLFP
GLIIYSDSLNHASMIEGIKRNAGPKRIFRHNDVAHLRELIAADDPAAPKLIAFESVYSMDGDFGPIKEICDIADEFG
ALTYIDEVHAVGMYGPRGAGVAERDGLMHRIDIFNGTLAKAYGVFGGYIAASAKMVDAVRSYAPGFIFSTSLPPAIA
AGAQASIAFLKTAEGQKLRDAQQMHAKVLKMRLKALGMPIIDHGSHIVPVVIGDPVHTKAVSDMLLSDYGVYVQPIN
FPTVPRGTERLRFTPSPVHDLKQIDGLVHAMDLLWARCALNRAEASA

Primers used: overhangs for ligation-independent cloning are shown in boldface.

RC_hemA_LIC_fwd – 5’-**TACTTCCAATCCAATGCA**GATTACAACTTGGCATTAGATAAAG-3’

RC_hemA_LIC_rev – 5’-**TTATCCACTTCCAATGCTA**AGCCGAGGCTTCTGCACG-3’

pET28b *Sphingomonas wittichii* serine palmitoyltransferase *E. coli*-optimized sequence. GenBank accession ABQ70245.1; linkers and the 6x-His tag are shown in bold.

atgggcagcagccatcatcatcatcacagcagcggcctggtgccgcgcggcagccatatggctagcatgactgg
tggacagcaaatgggtcgcggatccgaattcgagctcATGGCCGATTTGCTTAGCAAATTTGATCCTTTAATCGCCG
AGCGCGAGGCATTGCTTGCGACCGGCGTACGTGACCCGTACGCAATTGTTATGGACAAAGTGCTTTCGCCAACAGAA
GCCATGATTAATGGTCGTAAGACAATCTTGCTTGGTACGTATAATTACATGGGTATGACTTTCGATCCAGACGTGAT
TGCCGCCGGCAAGCAGGCGCTTGATGAATTTGGATCAGGAACAACAGGCTCCCGTGTCCTTAATGGAACTTATCAAG
GGCACAAAGCCTGCGAGGACGCACTGAAAGAGTTCTATGGAACAGAGCACGCTATCGTTTTCTCCACCGGATACCAG
GCTAACCTGGGTATGATCAGCACTCTGGCCGGTAAAGGTGACTATATCATTTTGGATGCCGACTCGCACGCCTCCAT
CTACGACGGTTGTTGGTTGGGAGACGCGGAGATCGTACGCTTCCGTCACAACAGTGTGGAGGACTTAGATAAGCGTT
TGGGTCGCCTTCCAGCCGAGGCCGGCAAATTAGTTGTTTTGGAGGGTGTCTACTCAATGATGGGCGACATCGCTCCC
CTGCAGGAGATGGTTGCAGTCTCAAAAAAACATGGGGCCATGATCTTAGTTGACGAAGCGCATGGGATGGGATTCTT
CGGGGAGCACGGTCGTGGTGTATTCGAAGAAGCAGGGGTAGAGGCAGACGTTGACTTCGTCGTTGGCACCTTTTCTA
AGTCAGTAGGGACGGTTGGTGGCTTCTGTGTGTCAAATCATCCCAAATTCGAAGTCCTTCGCTTGGTCTGCCGCCCT
TACGTCTTTACCGCATCATTGCCACCCAGCGTGGTAGCTACCGCGGCTACCAGCATCCGCAAGCTGATGCACGCAGG
CGACAAGCGCGCGCACTTATGGAAGAATAGTCGTCGTCTTCACCAAGGGTTGCGTGACATGGGCTACAAGTTGGGAA
CGGAGACAGCACAGTCTGCAATCATCGCTGTGATTTTGACGGACATGGCGCAGGCAGTAGCACTGTGGCAAGGGTTG
CTTGAGGCTGGGCTTTATGTAAATACGGCACGTCCGCCTGCTACGCCCGCCGGGATGTTCCTTTTGCGCTGCTCGTT
ATGTGCAGAACATAGTGATGAGCAAGTAGAGCAAATCCTGGGCATGTTTGAGTCGGCAGGGCGTGCTACGGGAGTAA
TTCCCAAGCTT**gcggccgcactcgagcaccaccaccaccaccac**

Translation: linkers and the 6x-His tag are shown in bold.

MGSSHHHHHHSSGLVPRGSHMASMTGGQQMGRGSEFELMADLLSKFDPLIAEREALLATGVRDPYAIVMDKVLSPTE
AMINGRKTILLGTYNYMGMTFDPDVIAAGKQALDEFGSGTTGSRVLNGTYQGHKACEDALKEFYGTEHAIVFSTGYQ
ANLGMISTLAGKGDYIILDADSHASIYDGCWLGDAEIVRFRHNSVEDLDKRLGRLPAEAGKLVVLEGVYSMMGDIAP
LQEMVAVSKKHGAMILVDEAHGMGFFGEHGRGVFEEAGVEADVDFVVGTFSKSVGTVGGFCVSNHPKFEVLRLVCRP
YVFTASLPPSVVATAATSIRKLMHAGDKRAHLWKNSRRLHQGLRDMGYKLGTETAQSAIIAVILTDMAQAVALWQGL
LEAGLYVNTARPPATPAGMFLLRCSLCAEHSDEQVEQILGMFESAGRATGVIPKL**AAALEHHHHHH**

C. Cloning

M. wollei sxtA ACP-AONS didomain (residues Ala722-Lys1243) as an *E. coli*-optimized sequence were ordered from Twist Biosciences subcloned into a pET28b plasmid (pET28b-*MW_sxtA_ACP-AONS*). The AONS domain (residues Glu822-Lys1243) was amplified in a 50 µL PCR reaction containing 10 µL Phusion HF buffer, 2 ng/µL parent plasmid, 2 µM each of forward and reverse LIC primers, 200 µM each of dNTPs, 0.04 U/µL Phusion HF. The didomain was amplified according to the following PCR procedure: 95 °C 2:00, (95 °C 0:30, 55 °C 1:00, 60 °C 2:00) for 30 cycles, 72 °C 10:00. The insert was purified by gel extraction and subcloned into pMCSG7K (custom plasmid provided by the University of Michigan Life Sciences Institute Center for Structural Biology) using standard LIC protocols,[74] resulting in pMCSG7K-*MW_sxtA_AONS*.

R. capsulatus hemA was obtained as an *E. coli*-optimized gene fragment from IDT and was amplified in a 50 µL PCR reaction containing 10 µL Phusion HF buffer, 0.8 ng/µL gene fragment, 2 µM each of forward and reverse LIC primers, 200 µM each of dNTPs, 0.04 U/µL Phusion HF. The didomain was amplified according to the following PCR procedure: 95 °C 2:00, (95 °C 0:30, 64 °C 0:45, 60 °C 1:30) for 25 cycles, 72 °C 10:00. The insert was purified by gel extraction and subcloned into pMCSG7 (provided by the University of Michigan Life Sciences Institute Center for Structural Biology) using standard LIC protocols,[74] resulting in pMCSG7-*RC_hemA*.

Correct insertion was confirmed by Sanger sequencing at the University of Michigan DNA Sequencing Core.

D. Protein overexpression and purification

Chemically competent BL21(DE3) cells were transformed with pMCSG7K-*MW_sxtA_AONS*. A single colony was picked to inoculate a 5 mL LB starter culture containing 50 µg/mL kanamycin and was incubated overnight at 37 °C, 200 rpm. The following day, 0.5 L TB media containing 50 µg/mL kanamycin was inoculated with the starter culture and incubated at 37 °C, 250 rpm until the OD_{600} reached 1.0. Cultures were equilibrated at 20 °C for 1 h prior to induction by addition of isopropyl-β-D-thiogalactopyranoside (IPTG, final concentration 200 µM). Cultures were incubated at 20 °C, 200 rpm for 18 h.

Harvested cell pellets were resuspended in approx. 4 mL of lysis buffer (50 mM HEPES, 300 mM NaCl, 10 mM imidazole, 1 mM PLP, 10% v/v glycerol, pH 8.0) per gram of wet pellet. The resuspended mixture was lysed by sonication (3 s on, 6 s off, 4:00 total). Insoluble material was removed by centrifugation (30,000 x *g* for 30 min at 4 °C). The clarified lysate was incubated with gentle shaking along with 2 mL of nickel-NTA resin for 1 h at 4 °C and poured into a column. The resin-bound protein was washed with 25 mL wash buffer (50 mM HEPES, 300 mM NaCl, 25 mM imidazole, 10% (v/v) glycerol, pH 8.0). The desired protein was eluted with 6 mL elution buffer (50 mM HEPES, 300 mM NaCl, 300 mM imidazole, 10% (v/v) glycerol, pH 8.0). The eluted protein was concentrated using 30 kDa centrifugal cutoff filters and exchanged into storage buffer (50 mM HEPES, 200 mM NaCl, pH 7.4) using a PD-10 column. The desalted protein was concentrated further using 30 kDa centrifugal cutoff filters, aliquoted, flash frozen in liquid nitrogen and stored at -80 °C. The typical yield was ~80 mg protein per half liter of TB culture.

HemA was expressed and isolated following same procedure but with the following buffers: lysis buffer contained 50 mM Tris-HCl, 300 mM NaCl, 10 mM imidazole, 1 mM TCEP, 1 mM PLP,

10% v/v glycerol, pH 8.0; wash elution buffer contained 50 mM Tris-HCl, 300 mM NaCl, 25 mM imidazole, 1 mM TCEP, 10% v/v glycerol, pH 8.0; buffer contained 50 mM Tris-HCl, 300 mM NaCl, 300 mM imidazole, 1 mM TCEP, 10% v/v glycerol, pH 8.0; storage buffer contained 50 mM HEPES, 200 mM NaCl, 10% v/v glycerol, pH 7.4.

SPT was expressed and isolated following same procedure but with the following buffers: lysis buffer contained 20 mM potassium phosphate, 150 mM NaCl, 10 mM imidazole, 1 mM PLP, 10% v/v glycerol, pH 7.5; wash buffer contained 20 mM potassium phosphate, 150 mM NaCl, 40 mM imidazole, 10% v/v glycerol, pH 7.5; elution buffer contained 20 mM potassium phosphate, 150 mM NaCl, 300 mM imidazole, 10% v/v glycerol, pH 7.5; storage buffer contained 50 mM HEPES, 200 mM NaCl, 10% v/v glycerol, pH 7.4.

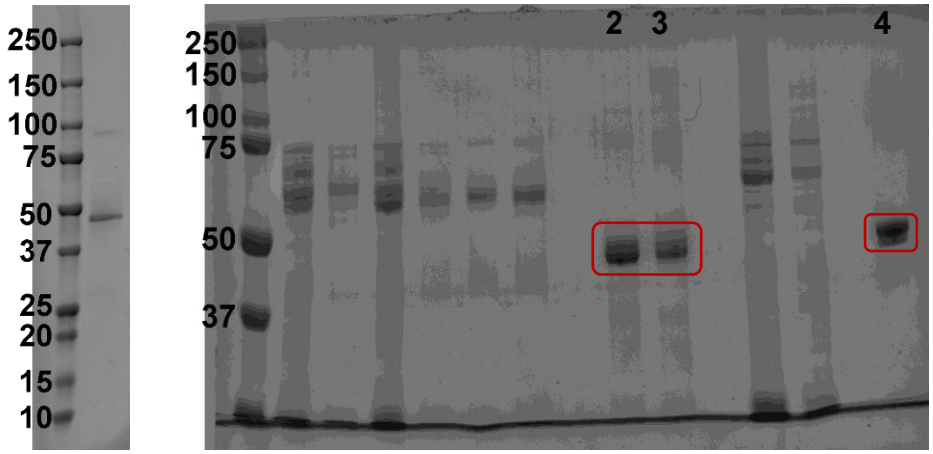


Figure 4.S1. Denatured SDS-PAGE protein gel of SxtA AONS, SPT and HemA (ALAS). Left: expected mass of recombinant *M. wollei* SxtA AONS = 49,641 Da. Right: lanes 2 and 3, expected mass of *S. wittichii* SPT = 48,549 Da; lane 4: expected mass of *R. capsulatus* HemA = 46,984 Da.

III. Enzymatic reactions

A. General information

Deuterium oxide was purchased from Cambridge Isotopes. Deuterated solutions were prepared by lyophilization on a Labconco FreeZone 2.5 L Freeze Dry System.

B. Preparation of deuterium-equilibrated solutions

Preparation of deuterated buffers

Reaction buffer was initially prepared in H_2O as a 10x concentrate (500 mM HEPES or potassium phosphate, 1.5 M NaCl, pH 7.0) and lyophilized. The remaining powder was resuspended in D_2O and lyophilized a second time. This powder was redissolved in D_2O to make a 10x stock solution.

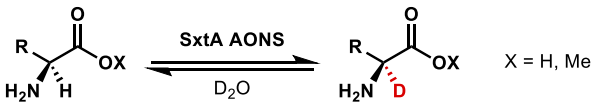
Preparation of deuterium-equilibrated protein

The 10x D_2O-equilibrated buffer above was diluted to the 1x working concentration (50 mM HEPES or phosphate, 150 mM NaCl) with additional D_2O. Protein aliquots of SxtA AONS were thawed and diluted with the 1x buffer, and then concentrated in 30 kDa MWCO centrifugal filters. The dilution/concentration cycle was repeated 3x total to exchange the protein into deuterated buffer. Protein concentration was measured again by the Pierce 660 assay and proteins were then used in assays on the same day.

Preparation of deuterium-equilibrated substrates

α-amino acids and their corresponding α-methyl esters were dissolved in D_2O to make 100 mM solutions. Less soluble acid substrates (Cys, Asp, Glu, Trp, Tyr) were prepared at 50 mM instead. The solutions were lyophilized, and the resulting powders or residues were redissolved in D_2O.

C. Enzymatic reaction procedures and data



Deuterium incorporation assays for mass spectrometry only (Figures 4.7, 4.8C, 4.10A, 13)

Buffer (50 mM HEPES, 150 mM NaCl diluted from 10x concentration above), 4 mM amino acid or methyl ester and 20 µM SxtA AONS in D_2O were combined for a total volume of 25 µL and incubated at 30 °C for 18 h. For PLP-only nonenzymatic control reactions, SxtA AONS was replaced with 20 µM PLP, and the substrate concentration varied between 4 and 20 mM (see Table 4.S2). For the timecourse reactions (Figure 4.8C), the total volume of the analytical-scale reactions was 100 µL. At every timepoint (1, 5, 10, 15, 20, 30, 60 and 120 min), 10 µL of the reaction mixture was withdrawn. In all cases, the enzymatic reactions were quenched with 4 volumes of MeOH. Precipitate was pelleted at 12,000 x *g* for 20 min. The supernatant was diluted 20x with acetonitrile and 1% formic acid in water for a final concentration of 4 µM substrate. The final ratios (v/v) for analysis were 5% quenched reaction, 49% of 1% formic acid in water and 45% acetonitrile. For reactions with higher concentrations of substrates, reaction aliquots were quenched as above but diluted further to a final substrate concentration of 4 µM such that the final LC-MS sample mixture was 50% aqueous and 50% organic by volume. Samples were analyzed by LC-MS: column = Waters Acquity 1.7 µm UPLC BEH Amide HILIC 2.1 x 100 mm; method = 85% B at 0.3 mL/min for 2 min, followed by a linear gradient to 75% B at 0.4 mL/min over 3 min and a second linear gradient to 60% B over 0.5 min, 60% B for 1 min and then 5.5 min re-equilibration at 85% B at 0.3 mL/min (total time 12 min). See Table 4.S1 for the retention times of each amino acid and methyl ester substrate tested.

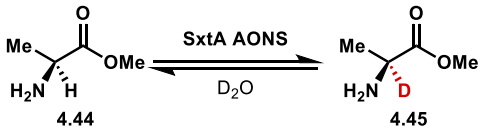
For early reactions with D_2O <99%, the stock buffer, substrate and protein solutions were used as normally prepared in H_2O and the remaining volume was filled with D_2O. Following reports that

blue light decreased the pKa of amino acid-PLP complexes,[75] we also attempted to irradiate translucent 96-well plate containing SxtA AONS and amino acids in D_2O with a blue Kessil lamp (450 nm). The irradiation suppressed almost all enzymatic activity.

Reaction with *C. raciborskii* SxtA ACP-AONS: Buffer (50 mM HEPES, 150 mM NaCl, 1 mM $MgCl_2$, 0.5 mM $MnCl_2$), 4 mM L-Arg and 10 µM SxtA ACP-AONS in D_2O were combined for a total volume of 25 µL and incubated at 30 °C for 18 h (76% v/v D_2O). Deuterium incorporation: 37%. Positive control with 1 mM propionyl-SPh: 81% deuteration.

Reactions with HemA/ALAS were performed with 50 mM HEPES pH 7.4, 150 mM NaCl, 1 mM $MgCl_2$, 4 mM L-Arg and 20 µM HemA in D_2O in a total volume of 25 µL and incubated at 30 °C for 18 h (76% v/v D_2O).

Reactions with SPT were performed with 20 mM potassium phosphate pH 7.4, 150 mM NaCl, 1 mM $MgCl_2$, 4 mM amino acids and 20 µM SPT in D_2O in a total volume of 25 µL and incubated at 30 °C for 18 h (81% v/v D_2O).



4.44 4.45

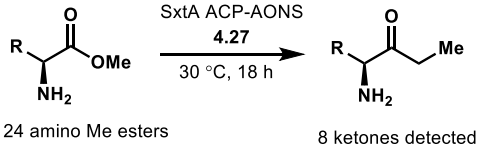
Recovery of deuterated L-Ala-OMe (4.45) on preparative scale (Figure 4.15)

Buffer (50 mM potassium phosphate, 150 mM NaCl diluted from 10x concentration above), 40 mM L-alanine methyl ester hydrochloride (**4.44**, 223 mg) and 20 µM SxtA AONS in D_2O were combined for a total volume of 40 mL and incubated at 30 °C for 8 h. The reaction was quenched by the addition of 160 mL of MeOH. Precipitate was pelleted at 5,000 x *g* for 20 min. The supernatant was concentrated to a white powder under reduced pressure. Ester **4.39** was dissolved in 4 mL of methanol and purified by flash chromatography on silica gel (10% v/v methanol in dichloromethane). The deuterated ester was isolated as the free base (98 mg, 60% recovery), with

99.2% deuterium incorporation. R_f = 0.2. **^{1}H NMR** (600 MHz, CD_3OD) δ 3.83 (s, 3H), 1.51 (s, 3H); **^{13}C NMR** (150 MHz, CD_3OD) δ 172.1, 53.6, 49.7 (t, *J* = 21 Hz), 16.5; **HRMS** (ESI) m/z calcd for $C_{18}H_{20}DFNO_2$ $[M+H]^+$ 105.0769 m/z, found 105.0773 (99.2% deuterium incorporation). 97.3% ee.

^{1}H NMR observation of deuterium incorporation into L-Arg-OMe (L-4.32, Figure 4.9)

Buffer (50 mM phosphate and 150 mM NaCl diluted from the 10x concentrated stock described above) and 20 mM L-Arg-OMe (**L-4.32**) were combined in D_2O in a microcentrifugue tube (450 µL final volume with enzyme). Upon addition of 20 µM SxtA AONS, the reaction mixture was immediately pipetted into an NMR tube and inserted into the NMR instrument (600 MHz) for observation. See Fig 4.9 for spectra.



Biosynthesis of α-amino ketones from α-amino methyl esters (Figure 4.11)

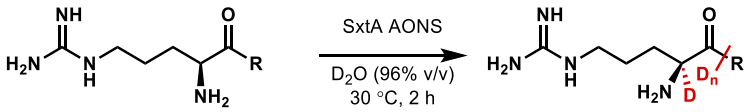
4 mM α-amino Me esters, 1 mM propionyl-pantetheine (**4.27**), and 20 µM *M. wollei* SxtA ACP-AONS were combined for a total reaction volume of 25 µL. The reactions were incubated at 30 °C for 18 h. Reactions were quenched by diluting 5x with methanol. Precipitate was pelleted at 12,000 x *g* for 20 min. The supernatant was diluted 20x with acetonitrile, 1% formic acid, and aqueous 10 µg/mL ^{15}N Arg·HCl. The final ratios (v/v) for analysis were 5% quenched reaction, 2.5% of aqueous 10 µg/mL ^{15}N Arg·HCl (Cambridge Isotope Laboratories), 46% acetonitrile, 46.5% of 1% formic acid in water (e.g., 10 µL of quenched reaction, 5 µL of 10 µg/mL ^{15}N Arg·HCl, 92 µL acetonitrile, 93 µL of 1% formic acid in water for a total volume of 200 µL).

Samples were analyzed by LC-MS: column = Waters Acquity 1.7 μm UPLC BEH Amide HILIC 2.1 x 100 mm column; method = 85% B at 0.3 mL/min for 2 min, followed by a linear gradient to 75% B at 0.4 mL/min over 3 min and a second linear gradient to 60% B over 0.5 min, 60% B for 1 min and then 6.5 min re-equilibration at 85% B at 0.3 mL/min (total time 15 min). t_R (^{15}N Arg): 4.0 min. Ratios were calculated by comparing the areas under the curve of the extracted ion chromatograms. The overall ketone yield was not calculated. Reactions were run in duplicate.

Table 4.S1. Summary of retention times of amino acid and methyl ester substrates by LC-MS

Side chain	Retention time of acid (min)	Retention time of methyl ester (min)
Gly	3.0	1.5
Ala	2.6	1.4
Val	2.0	1.1
Leu	1.6	1.1
Ile	1.7 (D-*allo*: 1.8 min)	1.1 (D-*allo*: 1.2)
Nle	1.6	1.1
Cys	2.8	1.1
Met	2.0	1.1
Asn	3.8[a]	1.8
Gln	3.6[a]	1.7
Ser	3.7	1.7
Thr	3.2 (D-*allo*: 3.3)	1.5 (D-*allo*: 1.7)
Phe	1.7	1.1
Tyr	2.1	1.2
Trp	1.5	1.1
His	4.4	2.8
Lys	4.4	2.8
Orn	4.6	3.0[b]
Arg	4.3	2.5
Canavanine	4.3	2.7
Homoarginine	4.0	2.4
Citrulline	3.7	1.9
Asp	3.4	1.8
Glu	3.0	1.5

[a]The masses corresponding to the free acids were only detected in reactions with the methyl esters substrates and not in reactions with the amino acid substrates. [b]A portion also appears as the cyclized adduct.



Deuterium incorporation into arginine ketone standards

Buffer (50 mM potassium phosphate, 150 mM NaCl diluted from 10x concentration above), 2 mM synthetic ketone standards (synthesized by Dr. Meagan Hinze, see Chapter 3.4) and 20 µM SxtA AONS in D_2O were combined for a total volume of 25 µL (96% v/v) and incubated at 30 °C for 2 h. The enzymatic reactions were quenched with 4 volumes of MeOH. Precipitate was pelleted at 12,000 x *g* for 20 min. The supernatant was diluted 20x with acetonitrile and 1% formic acid in water. Samples were analyzed by LC-MS: column = Waters Acquity 1.7 µm UPLC BEH Amide HILIC 2.1 x 100 mm; method = 85% B at 0.3 mL/min for 2 min, followed by a linear gradient to 75% B at 0.4 mL/min over 3 min and a second linear gradient to 60% B over 0.5 min, 60% B for 1 min and then 5.5 min re-equilibration at 85% B at 0.3 mL/min (total time 12 min).

Because of the additional acidic α-protons on the ketones, they may be deuterated more than once.

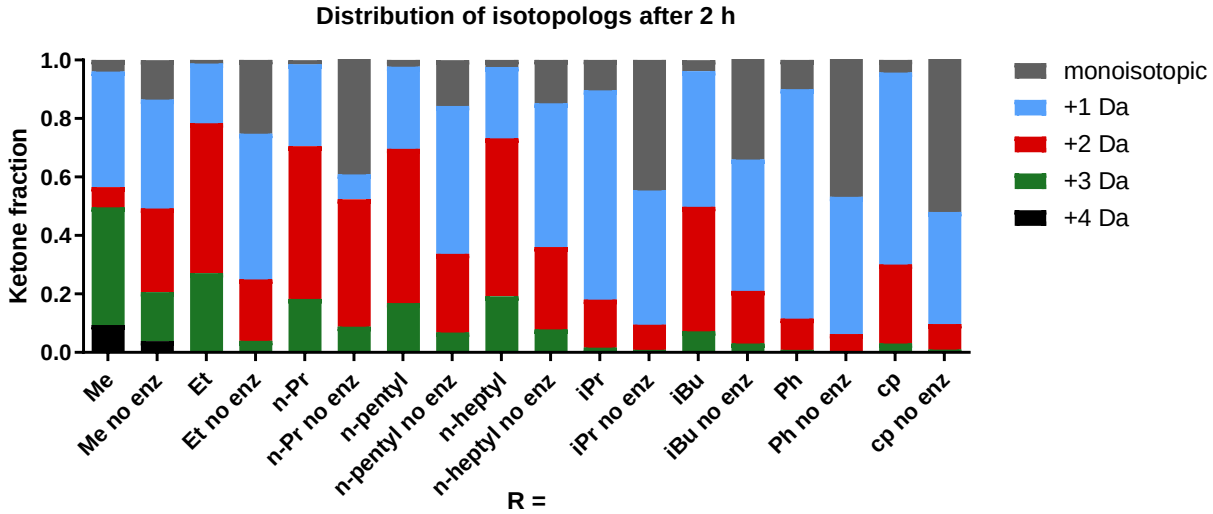


Figure 4.S2. High deuterium incorporation into synthetic arginine-based ketones.

Table 4.S2. Deuterium incorporation into amino acid and methyl ester substrates (4 mM vs. 20 mM)

Side chain	4 mM			20 mM		
	L- amino acid	L- methyl ester	D- methyl ester	L- methyl ester	D-methyl ester	PLP only with D- Me esters
Gly	11.8% ± 0.6	98.2% ± 0.3	N/A	99.7% ± 0.1	N/A	6.2% ± 0.1
Ala	18.8% ± 1.3	97.9% ± 0.1	17.0% ± 1.8	99.6% ± 0.1	14.0% ± 0.1	1.0% ± 0.1
Val	11.9% ± 0.6	97.3% ± 0.5	2.7% ± 0.1	99.6% ± 0.1	0.7% ± 0.1	0.5% ± 0.1
Leu	8.1% ± 1.3	67.7% ± 0.1	21.1% ± 1.4	22.6% ± 2.1	11.2% ± 0.2	0.8% ± 0.1
Ile	4.5% ± 0.3	97.4% ± 0.4	1.5% ± 0.1	98.4% ± 0.1	0.4% ± 0.1	0.7% ± 0.1
Nle	35.0% ± 0.1	97.9% ± 0.1	26.8% ± 0.3	96.8% ± 0.4	26.3% ± 0.1	1.1% ± 0.1
Cys	85.3% ± 0.2	37.2% ± 2.0	6.2% ± 0.6	2.6% ± 0.1	1.5% ± 0.1	5.7% ± 0.1[a]
Met	23.9% ± 0.2	97.9% ± 0.1	45.8% 0.3	39.0% ± 0.1	18.3% ± 0.2	1.0% ± 0.1
Asn	ND	94.4% ± 0.5	68.6% ± 0.4	94.4% ± 1.0	48.6% ± 3.1	11.9% ± 0.7
Gln	ND	97.5% ± 0.1	6.3% ± 0.7	71.3% ± 0.2	3.8% ± 0.4	2.8% ± 0.7
Ser	37.2% ± 0.6	97.7% ± 0.4	68.2% ± 0.8	99.6% ± 0.1	61.7% ± 0.5	3.1% ± 0.1
Thr	64.3% ± 0.1	98.2% ± 0.1	98.2% ± 0.1	99.6% ± 0.1	5.0% ± 0.5	1.9% ± 0.1
Phe	11.6% ± 0.3	97.4% ± 0.5	82.2% ± 0.7	69.2% ± 0.1	38.4% ± 0.5	7.1% ± 0.1
Tyr	14.2% ± 0.5	98.7% ± 0.1	98.5% ± 0.1	99.8% ± 0.1	92.3% ± 0.4	2.6% ± 0.1
Trp	12.6% ± 1.1	97.7% ± 0.1	54.0% ± 0.3	99.7%[b]	22.6% ± 0.1	0.2% ± 0.1
His	97.3% ± 1.2	96.2% ± 0.7	71.3% ± 0.8	0%	0%	7.0% ± 0.2[a]
Lys	0%	98.0% ± 0.1	4.8% ± 0.1	85.8% ± 2.5	0%	4.6% ± 0.1[a]
Orn	0%	94.4% ± 0.1	8.9% ± 0.1	7.0%[b]	0%	3.6% ± 0.4[a]
Arg	37.4% ± 0.7	98.4% ± 0.3	97.3% ± 0.7	99.6% ± 0.1	94.3% ± 0.6	3.0% ± 0.1[a]
Canavanine	97.9% ± 0.1	96.7% ± 0.1	N/A	99.3% ± 0.1	N/A	N/A
Homoarg.	0.6% ± 0.1	97.6% ± 0.3	26.6% ± 3.1	97.4% ± 0.4	1.4%[b]	2.3% ± 0.2[a]
Citrulline	0%	96.5% ± 0.6	11.9% ± 0.1	0%	0%	3.6% ± 0.1[a]
Asp	40.7% ± 2.8	36.9% ± 0.3	1.2% ± 0.1	22.7% ± 0.3	0.2% ± 0.1	1.7% ± 0.1
Glu	0%	27.2% ± 0.3	3.9% ± 0.7	14.5% ± 0.1	2.2% ± 0.1	12.7% ± 0.1

[a]Used 4 mM ester substrate. [b]One sample was analyzed.

Table 4.S3. Deuterium incorporation in nonenzymatic control reactions after 18 h (omitted from Figure 4.8C)

Substrate	Additive	Deuterium incorporation
L-Arg (**L-4.19**)	20 μM PLP	1.2% ± 0.1
L-Arg (**L-4.19**)	none	0.9% ± 0.1
L-Arg-OMe (**L-4.32**)	20 μM PLP	9.3% ± 0.1
L-Arg-OMe (**L-4.32**)	none	0.7% ± 0.1
Product ketone standard **4.31**	20 μM PLP	42.9% ± 0.3
Product ketone standard **4.31**	none	28.9% ± 0.1

Chromatograms were extracted for the monoisotopic mass, *m*+1 and *m*+2 (one and two heavy atoms, respectively) for *z* = 1, and integrated. The natural abundances of heavy atom isotopologs (^{13}C, ^{15}N, ^{34}S, as calculated by enviPat https://www.envipat.eawag.ch/index.php) were subtracted from the *m*+1 and *m*+2 integration values before calculating the fraction of deuterated amino acid or methyl ester. A small level of nonenzymatic deuterium incorporated was observed in a few substrates, which were subtracted from the enzymatic reaction totals. For methyl ester samples, the acids resulting from hydrolyzed esters were not included in the deuteration calculation. Deuterium incorporation values listed above are the average of two samples.

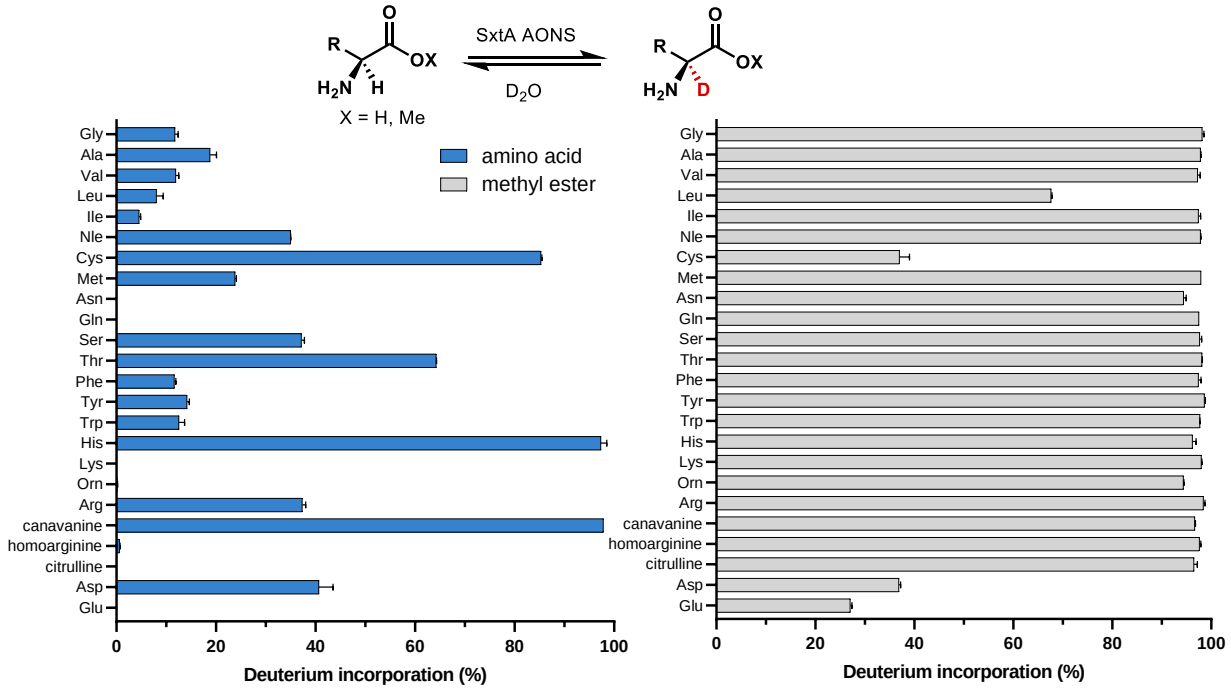


Figure 4.S3. Plots of α-deuteration after esterification of L substrates at 4 mM after overnight incubation (see Table 4.S2).

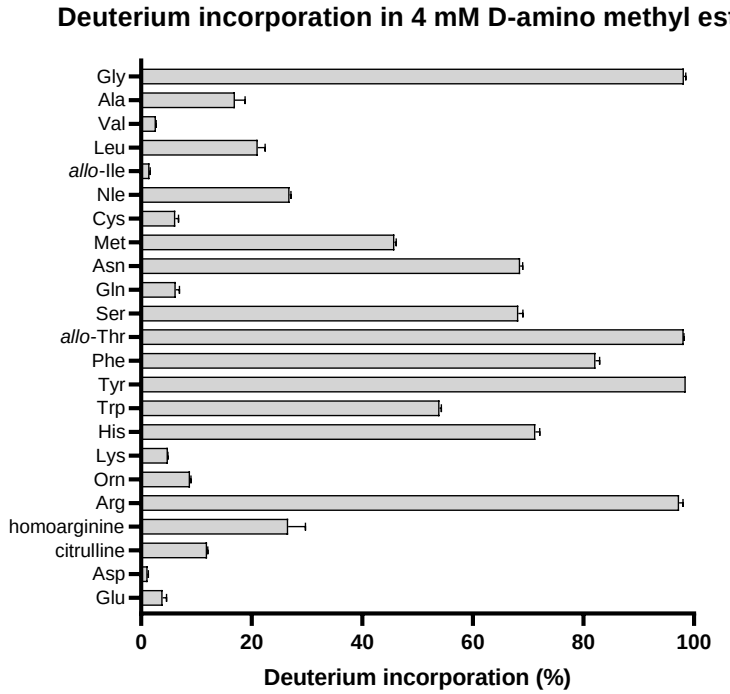


Figure 4.S4. SxtA AONS-mediated deuterium incorporation in D-amino methyl esters only, overnight incubation (see Table 4.S2).

D. Determination of enantioenrichment

The amine groups of amino acid and methyl ester substrates were Fmoc-protected prior to chiral supercritical fluid chromatography analysis.

Preparation of enantiopure and racemic standards for supercritical fluid chromatography

Adapted from Zhou et al.[76] Buffer (50 mM HEPES pH 7.0, 150 mM NaCl), 20 mM total methyl ester, and 0-40 mM NaOH (an equimolar amount to neutralize HCl salts originating from the stock solutions, e.g., L-Arg-OMe·2HCl, see Table 4.S4) were combined for a total volume of 50 µL. This buffer solution was diluted with 100 µL of MeCN and basified slightly with 50 µL of borate buffer (40 mM, pH 9.2), and then derivatized with 50 µL of an Fmoc-Cl solution (80 mM in acetonitrile). After 15 min at rt, reactions were centrifuged at 10,000 x *g* for 10 min and 200 µL of the supernatant was transferred to sample vials for SFC analysis.

Table 4.S4 Example derivatization for the Arg-OMe racemic standard

Stock solution	Volume
1 M HEPES buffer pH 7.0	5 µL
3 M NaCl	5 µL
100 mM L-Arg-OMe·2HCl	10 µL
100 mM D-Arg-OMe·2HCl	10 µL
1 M NaOH	4 µL
water	16 µL
MeCN	100 µL
40 mM borate buffer pH 9.2	50 µL
80 mM Fmoc-Cl in MeCN	50 µL
Total volume	250 µL

Deuterium incorporation for methyl ester substrates and preparation of reaction mixtures for SFC and MS (Figures 4.8D and 4.10B)

Enzymatic reactions in duplicate were performed with 50 mM HEPES original pH 7, 150 mM NaCl, 4 or 20 mM methyl ester (varies by substrate, see Table 4.S8), and 20 µM SxtA AONS in D_2O (final volume scaled by the number of desired timepoints). For every timepoint sample, 100 µL (for SFC analysis) and 10 µL (for LC-MS analysis) aliquots were removed and stopped by flash freezing in liquid nitrogen. Thawed aliquots for LC-MS analysis were prepared as described above in subsection "Deuterium incorporation assays for mass spectrometry only". Thawed SFC aliquots were quenched with 4 volumes of MeOH and precipitate was pelleted at 12,000 x *g* for 20 min. The supernatant was transferred to a glass test tube and all liquids were removed under reduced pressure. The residues were redissolved in water and NaOH to neutralize any HCl salts for a total volume of 50 µL. This buffer solution was diluted with 100 µL of MeCN and basified slightly with 50 µL of borate buffer (40 mM, pH 9.2), and then derivatized with 50 µL of an 80 mM Fmoc-Cl solution in acetonitrile. After 15 min at rt, reactions were centrifuged at 10,000 x *g* for 10 min and 200 µL of the supernatant was transferred to sample vials for SFC analysis.

For reactions at 4 mM substrate concentration (Orn), the supernatant was diluted 1:1 v/v with MeOH before transferring to sample vials.

Table 4.S5. Example SFC preparation for a racemic Arg-OMe reaction originally with 20 mM total methyl esters

Stock solution	Volume
Quenched reaction residue	-
1 M NaOH	4 µL
water	46 µL
MeCN	100 µL
40 mM borate buffer pH 9.2	50 µL
80 mM Fmoc-Cl in MeCN	50 µL
Total volume	250 µL

Chiral separation by SFC: Co-solvent = MeOH with 0.5% v/v formic acid. Column = Daicel CHIRALCEL 5 µm OD-H 1.6 x 250 mm. Isomers were separated with the isocratic methods listed in Table 4.S7. Residual Fmoc-Cl, hydrolyzed Fmoc-OH and derivatized HEPES were also visible in some methods. The UV signal of Fmoc derivatives was monitored at 262 nm. Safinamide samples were monitored at 228 nm but could not be resolved. Deuterium incorporation and er/dr values in Table 4.S8 are an average of two samples. Background activity seen in the nonenzymatic negative controls and contaminants in single stereoisomer standards were subtracted.

Table 4.S6. Deuterium incorporation and stereoretention in Figure 4.8D

Time	L-Arg-OMe		D-Arg-OMe		Racemic Arg-OMe	
	Deuterium	Stereoretention	Deuterium	Stereoretention	Deuterium	er (D:L)
0 min	0%	100%	0%	100%	0%	50:50
5 min	42.3%	98.8%	-	-	-	-
10 min	60.4%	98.8%	-	-	-	-
15 min	74.3%	98.6%	-	-	43.5%	49:51
30 min	93.1%	98.5%	-	-	55.9%	49:51
45 min	96.5%	94.7%	-	-	61.2%	47:53
1 h	97.1%	97.2%	31.6%	96.3%	65.1%	47:53
2 h	97.1%	94.9%	52.8%	79.8%	75.4%	44:56
3 h	-	-	66.0%	81.0%	80.9%	44:56
4 h	-	-	72.1%	77.5%	83.1%	43:57
5 h	-	-	77.9%	73.2%	85.1%	44:56
6 h	96.8%	90.8%	81.6%	69.9%	86.2%	43:57
18 h	96.6%	90.1%	88.9%	61.8%	87.7%	43:57

Table 4.S7. Separation conditions and retention times of α-amino methyl esters by SFC[a]

Side chain	Cosolvent concentration	t_R Fmoc-OH (min)	t_R D isomer (min)	t_R L isomer (min)	Notes:
Gly	N/A	N/A	N/A	N/A	The configuration of singly deuterated species cannot be determined on this instrument.
Ala	15%	6.4	5.8	8.3	
Val	15%	6.4	4.4	5.3	
Leu	15%	6.4	4.7	5.3	
Ile	15%	6.4	4.5	5.2	
Nle	12%	8.4	7.0	8.4	See footnote b
Cys	25%	3.7	5.0	6.4	See footnote c
Met	25%	3.7	4.6	5.2	
Asn	25%	3.7	6.5	7.8	
Gln	30%	3.1	5.2	5.8	
Ser	30%	3.7	4.3	6.0	
Thr	10%	10.2	16.0	17.4	
Phe	25%	3.7	5.8	7.3	
Tyr	30%	3.1	5.3	6.9	
Trp	35%	2.7	10.9	11.7	
His	-	-	-	-	See footnote d
Lys	-	-	-	-	See footnote d
Orn	40%	2.7	5.1	8.4	
Arg	31%	3.0	4.1	5.3	
Canavanine	-	-	-	-	See footnote d
Homoarginine	-	-	-	-	See footnote d
Citrulline	-	-	-	-	See footnote d
Asp	-	-	-	-	See footnote d
Glu	-	-	-	-	See footnote d
Safinamide	-	N/A	-	-	See footnote d

[a]At 25% cosolvent and above, a blank MeOH injection is required between each sample to clear derivatized HEPES. [b]The L isomer and Fmoc-OH co-elute, so only 0.5 equiv of Fmoc-Cl was added to the derivatization reaction. The ratios may still be skewed, so these data points were not included in Figure 4.10B. [c]Required the addition of 100 mM TCEP to the sample mixture to reduce the cystine disulfides back to cysteine methyl ester. [d]The most polar side chains and safinamide could not be separated by SFC.

Table 4.S8. Enantiomeric and diastereomeric ratios at the time of highest deuterium incorporation in Figure 4.10B

Side chain	L ester deuterium incorporation	L ester stereoretention	Time (h)	D ester deuterium incorporation	D ester stereoretention	Time (h)
Glycine	99.7%	N/A	18[a]	N/A	N/A	N/A
Alanine	99.6%	99.3%	4[a]	14.0%	88.6%	18[a]
Valine	99.5%	99.7%	8[a]	0.7%	100%	18[a]
Leucine	22.6%	98.2%	18[a]	11.2%	92.4%	18[a]
Isoleucine	98.4%	100%	4[a]	0.4%	100%	18[a]
Norleucine	96.8%	99%[c]	18[a]	26.3%	88.3%[c]	18[a]
Cysteine	2.6%	98.3%	8[a]	1.5%	100%	18[a]
Methionine	39.0%	99.5%	18[a]	18.3%	81.5%	18[a]
Asparagine	94.5%	97.5%	8[a]	48.6%	75.8%	18[a]
Glutamine	71.3%	100%	8[a]	3.8%	100%	18[a]
Serine	99.6%	90.4%	8[a]	61.7%	66.6%	18[a]
Threonine	99.6%	98.6%	18[a]	5.0%	97.3%	18[a]
Phenylalanine	69.2%	98.4%	18[a]	38.4%	76%	18[a]
Tyrosine	99.8%	59.3%	8[a]	92.3%	47%	18[a]
Tryptophan	99.7%	98.4%	8[a]	22.6%	86.9%	18[a]
Histidine	97.9%	-	18[b]	71.4%	-	18[b]
Lysine	85.8%	-	18[a]	4.8%	-	18[b]
Ornithine	94.4%	99.5%	18[b]	8.5%	89.4%	18[b]
Arginine	99.6%	96%	1[a]	94.3%	43.8%	18[a]
Canavanine	99.3%	-	2[a]	N/A	-	N/A
Homoarginine	97.4%	-	6[a]	17.2%	-	18[b]
Citrulline	96.5%	-	18[b]	11.9%	-	18[b]
Aspartic acid	22.8%	-	18[a]	0.2%	-	18[a]
Glutamic acid	14.5%	-	18[a]	2.2%	-	18[a]

The enantiomers of safinamide, and the most acidic and basic side chains could not be resolved well by SFC. [a]20 mM ester. [b]4 mM ester. [c]Values may be skewed by the co-elution of the Fmoc-OH and L-Nle-OMe. These were omitted from Figure 4.10B but are included in Figure 4.S5 and Figure 4.S6 below.

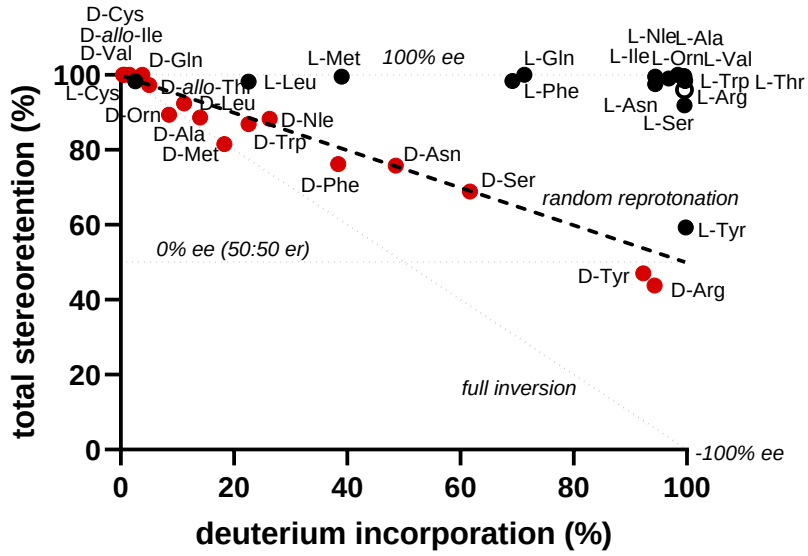


Figure 4.S5. Plot of endpoint deuterium incorporation vs. stereoretention in Table 4.S8. Norleucine and cysteine data are included.

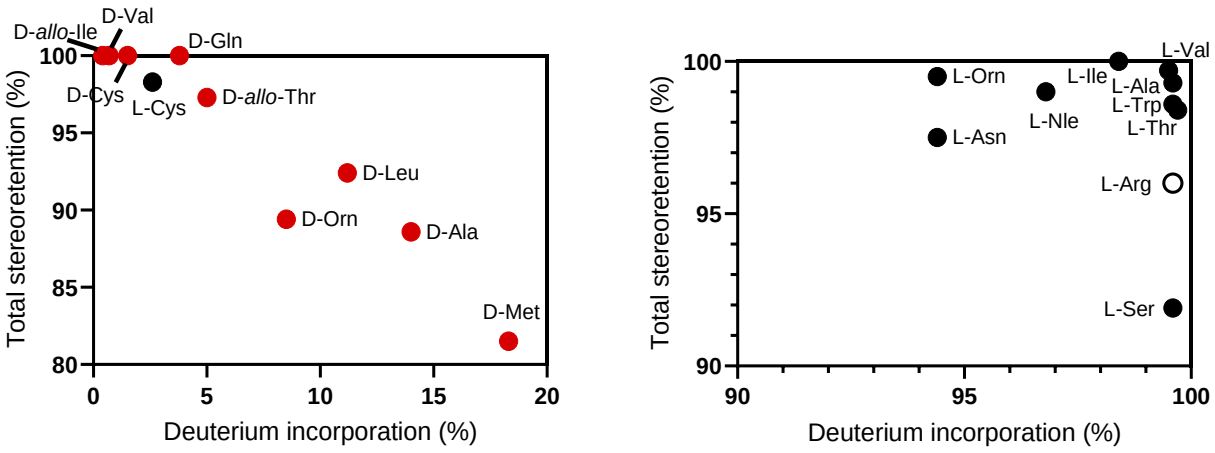
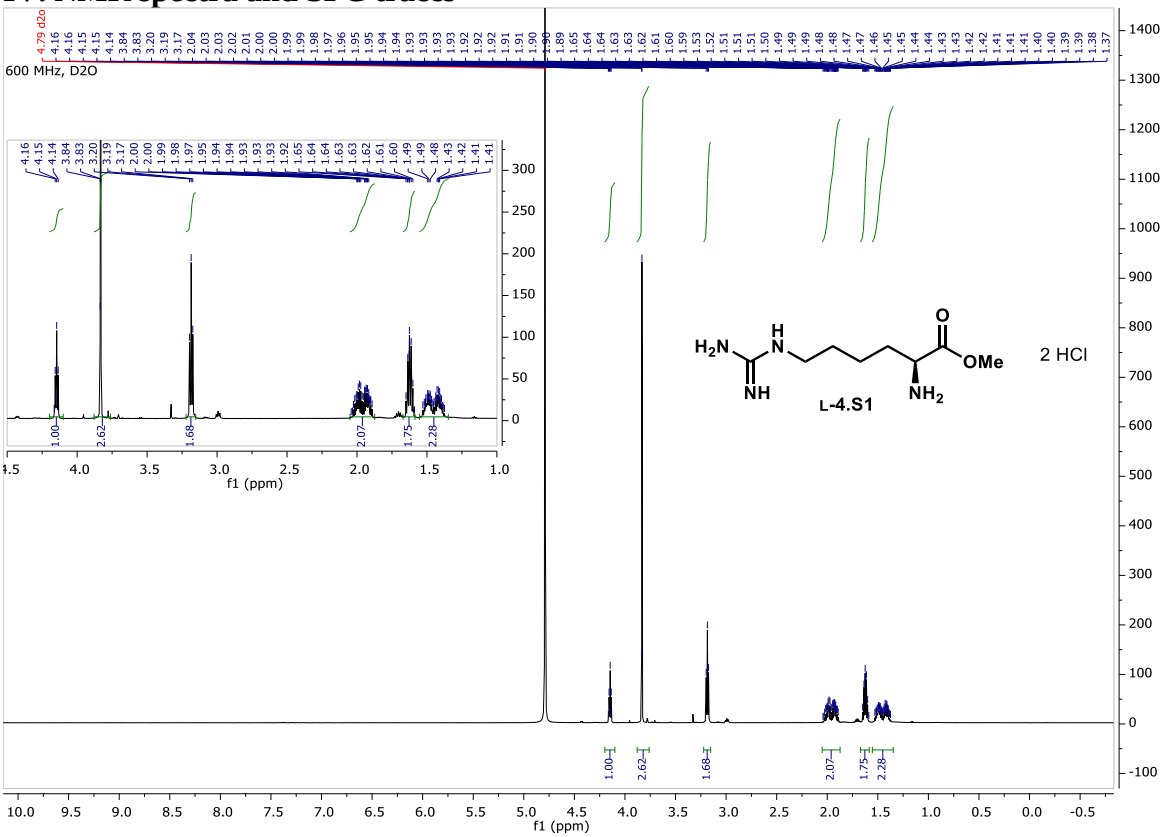
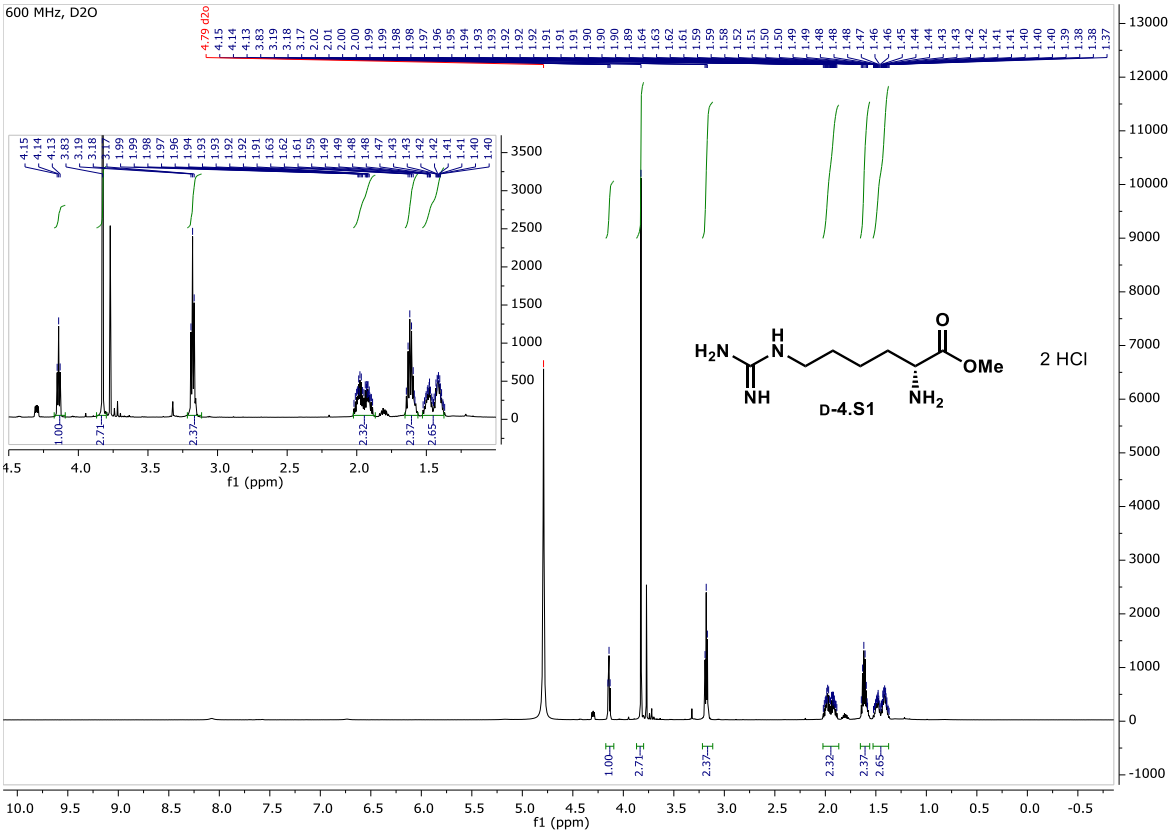
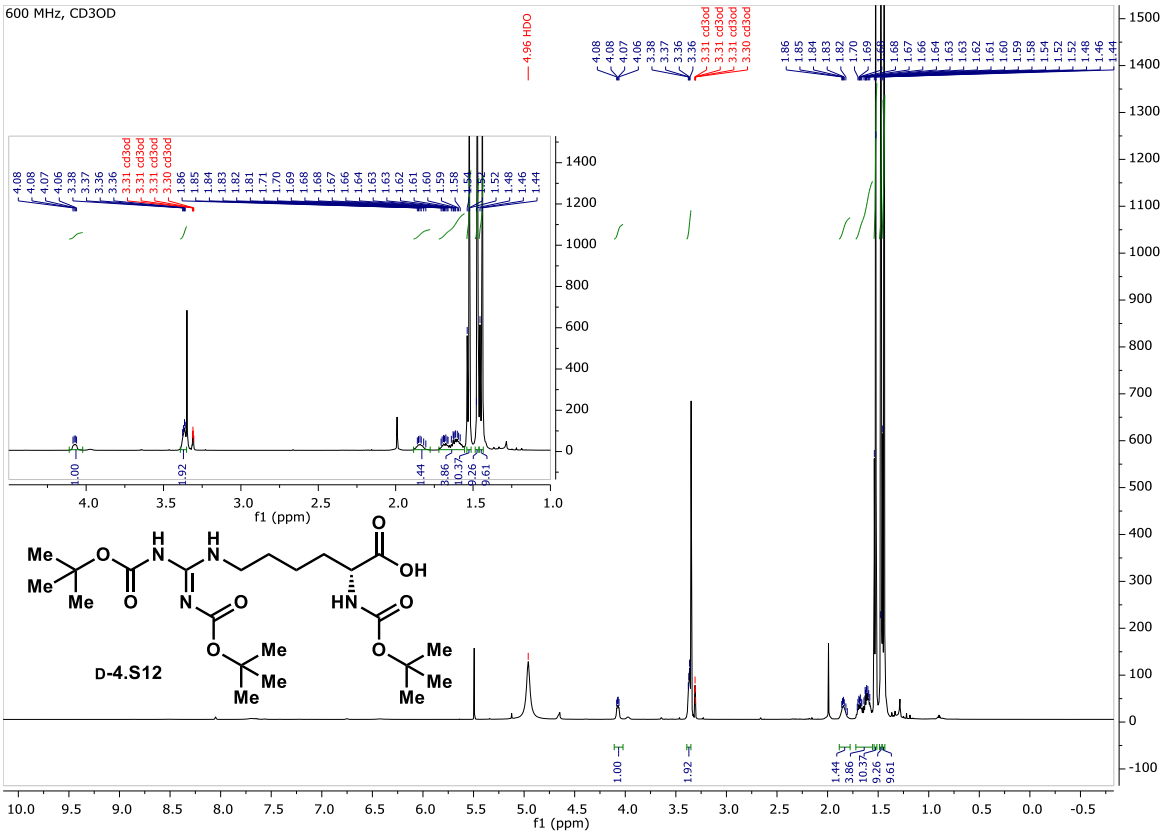


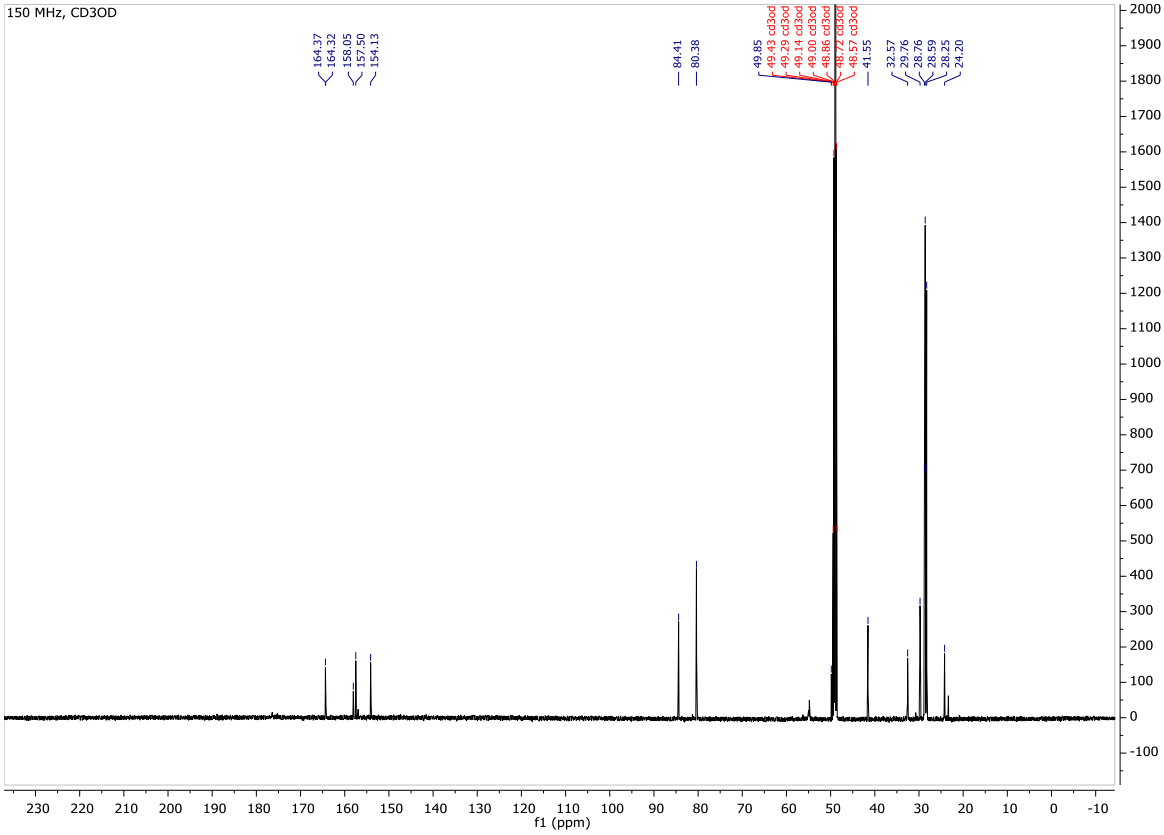
Figure 4.S6. Expansions of crowded clusters of Figure 4.S5.
Top left corner of Figure 4.S5 (left). Top right corner of Figure 4.S5 (right).

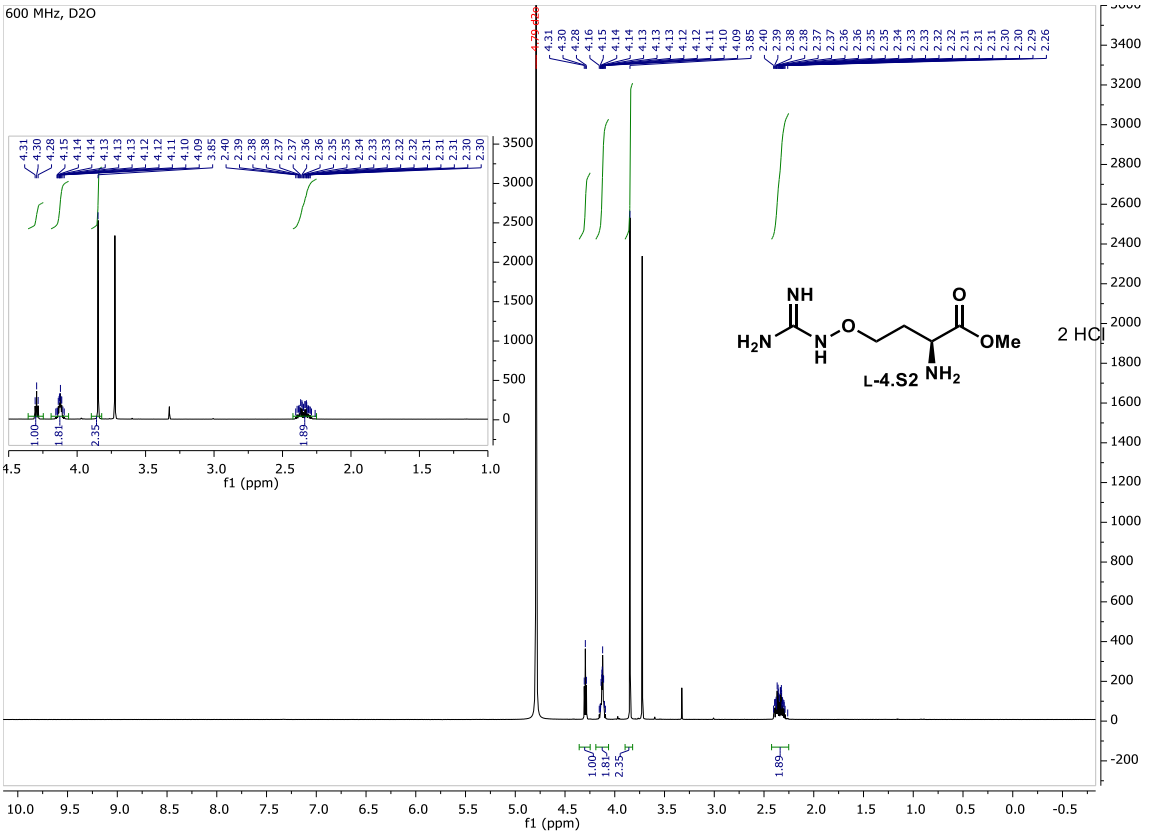
IV. NMR spectra and SFC traces

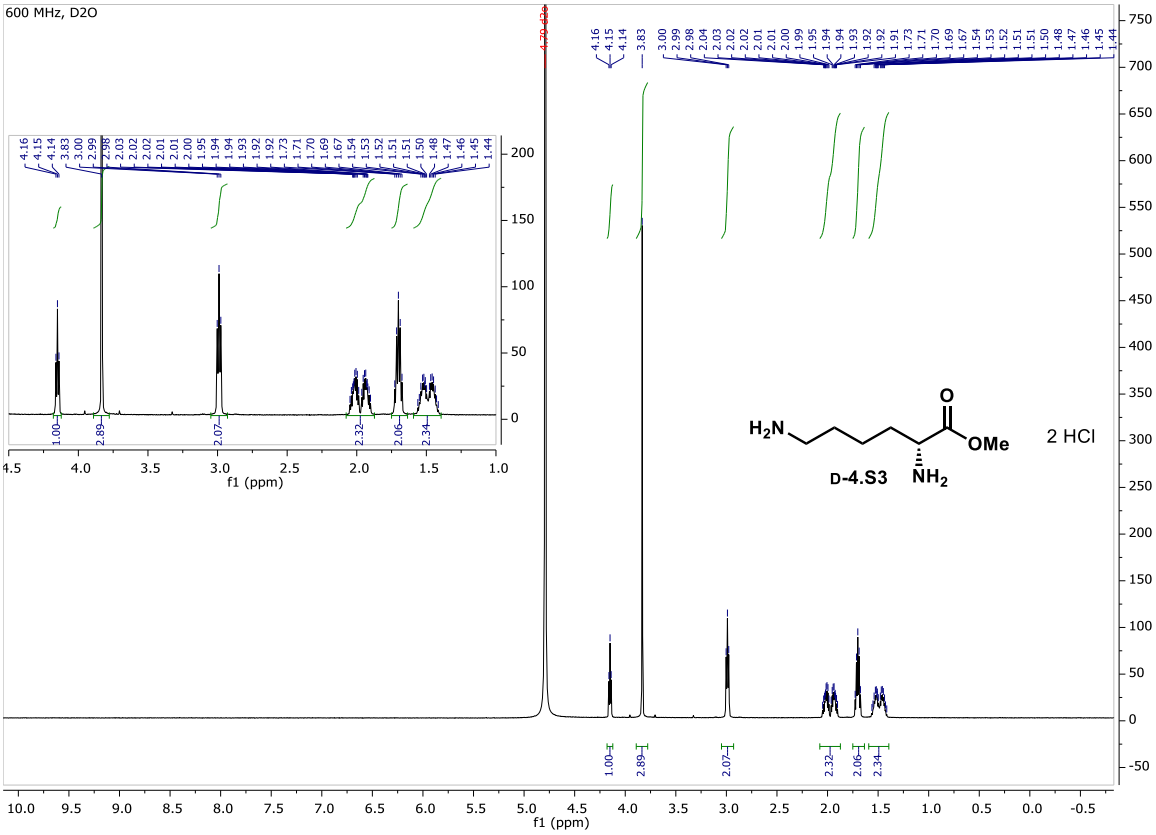


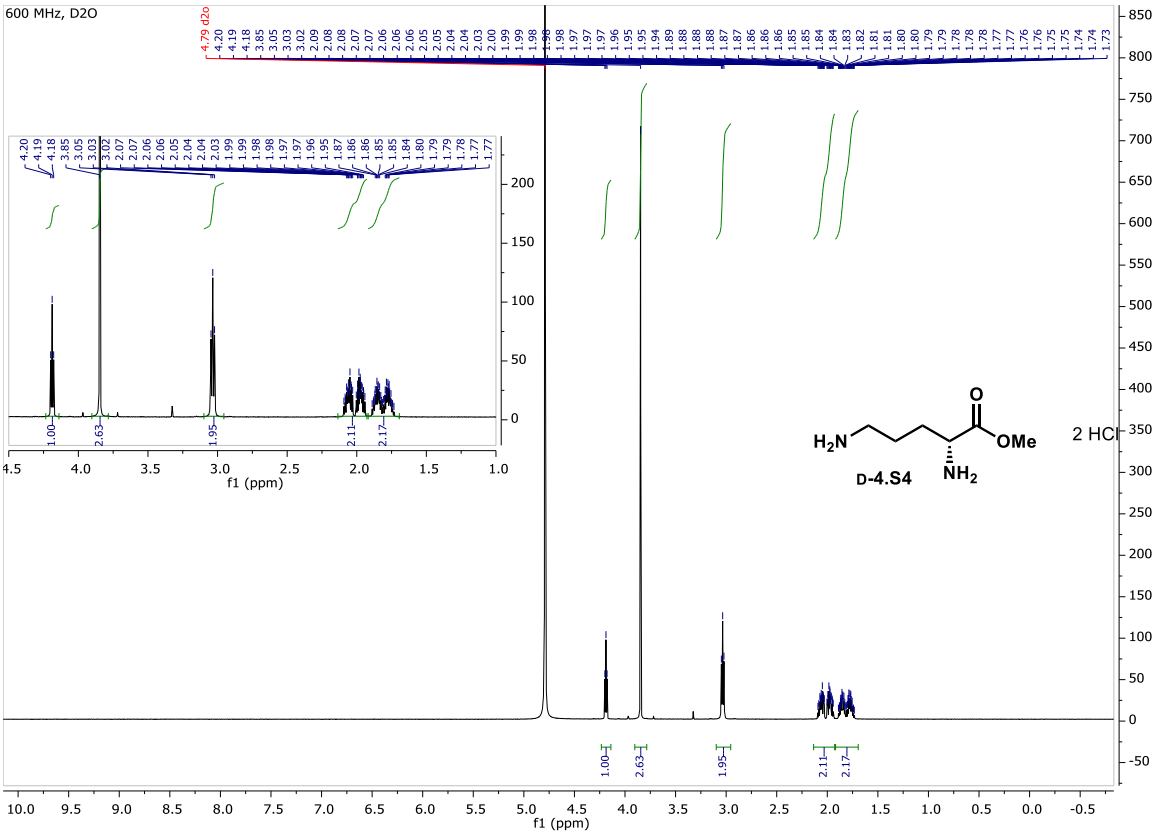


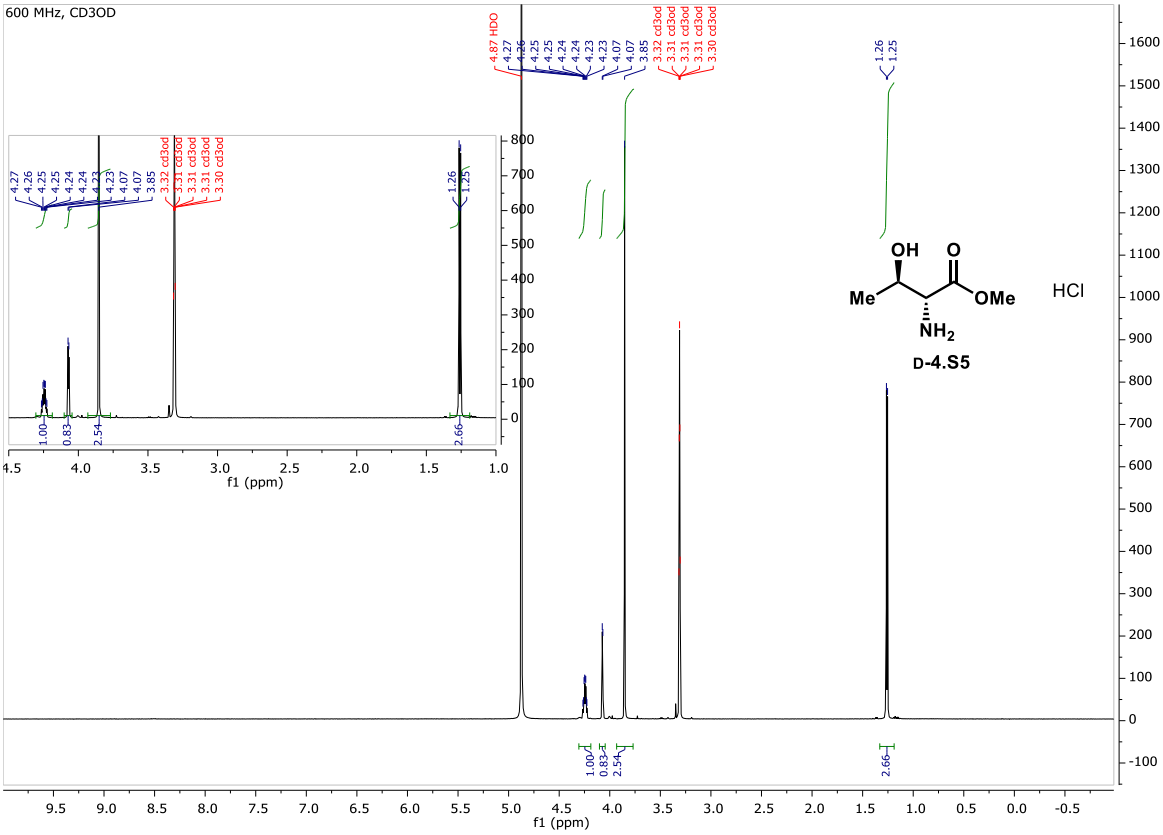

600 MHz, CD3OD
D-4.S12
f1 (ppm)

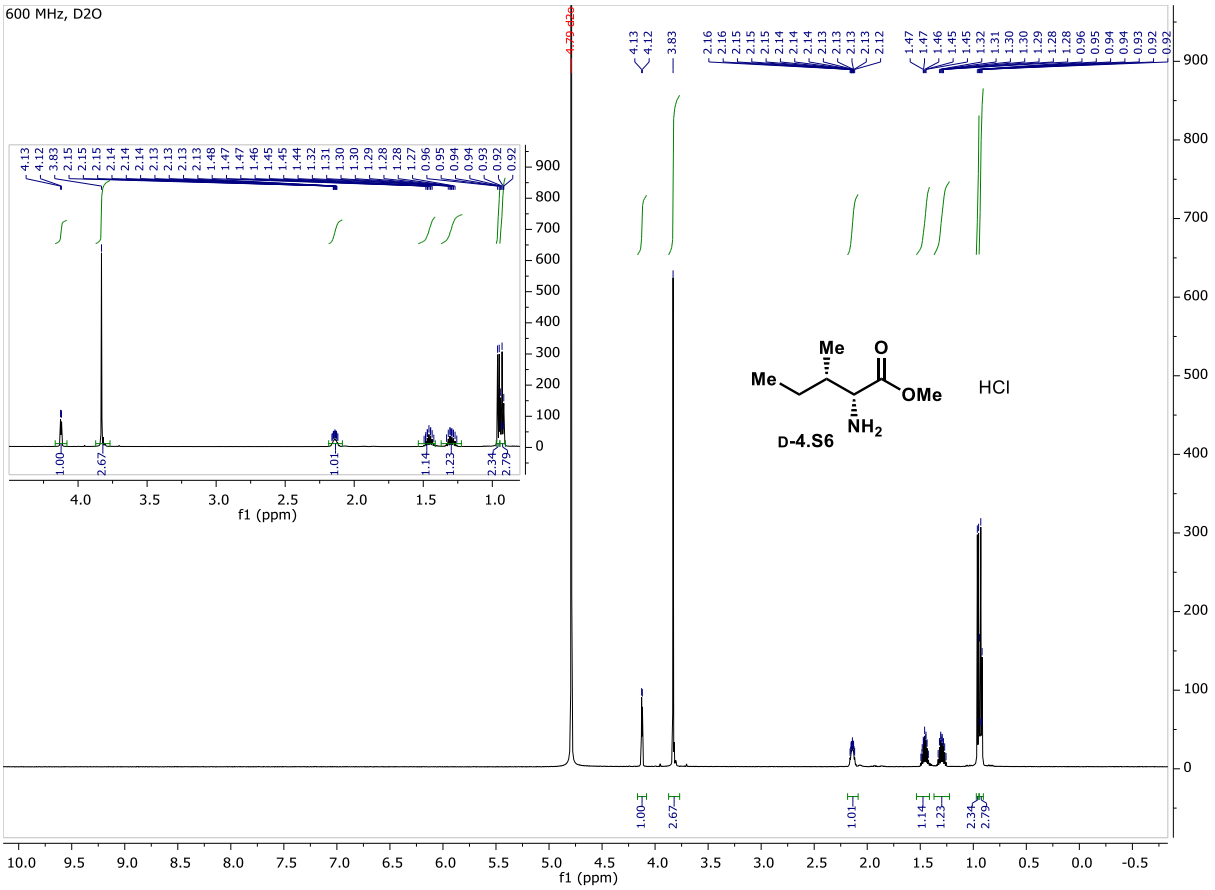

150 MHz, CD3OD
f1 (ppm)

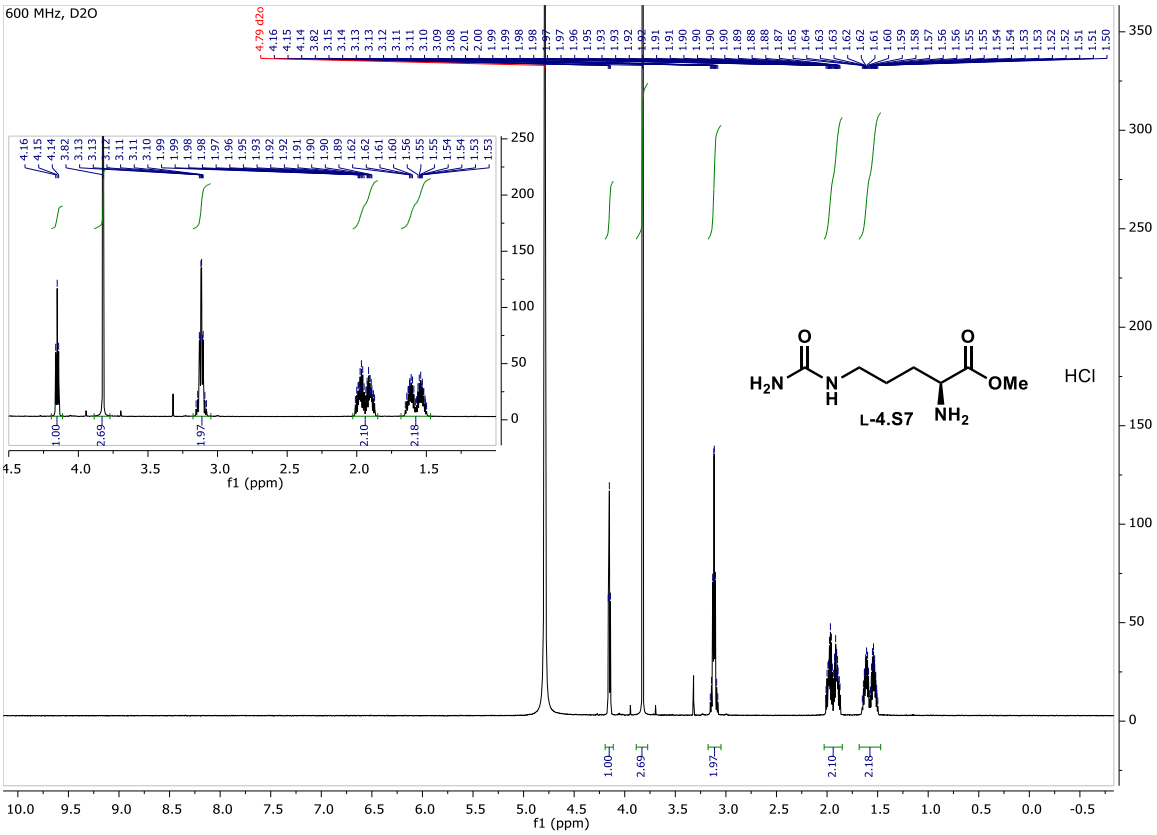

600 MHz, D2O
NH
O
H2N
N
H
O
OMe
NH2
L-4.S2
2 HCl
f1 (ppm)

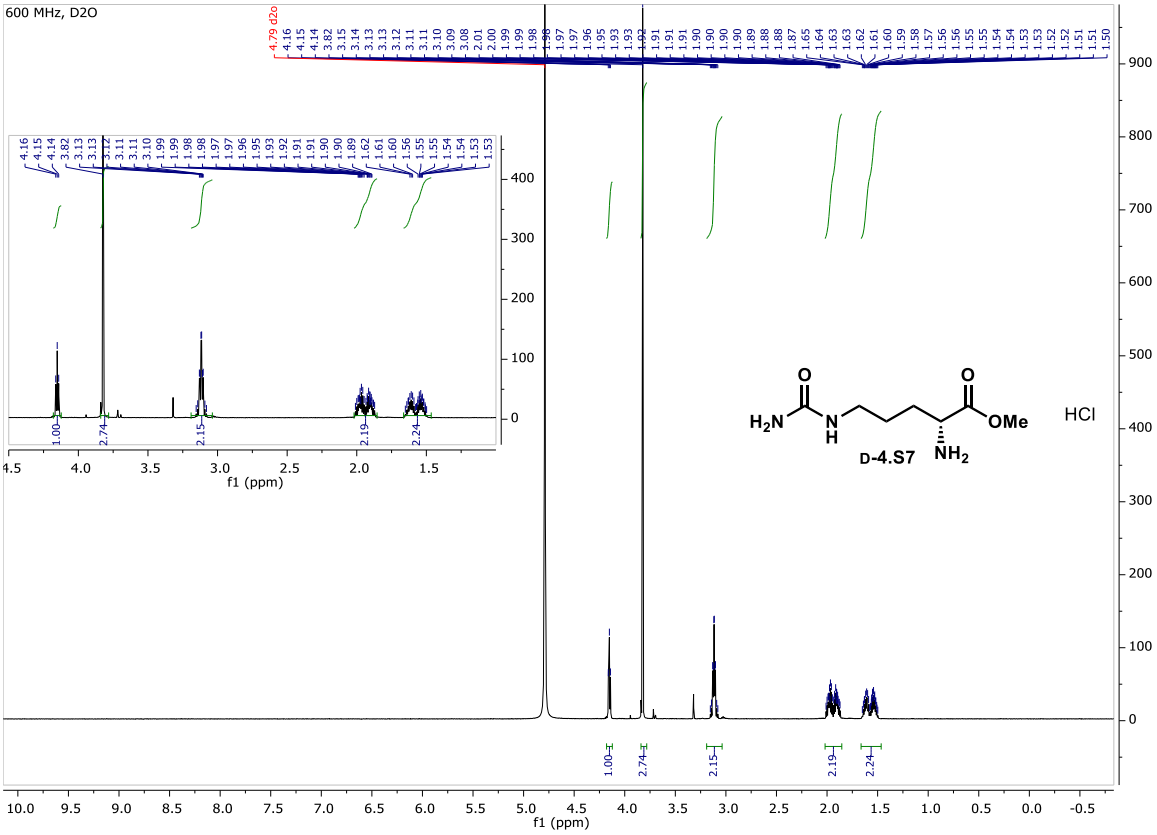

600 MHz, D2O
O
H2N
OMe
NH2
D-4.S3
2 HCl
f1 (ppm)

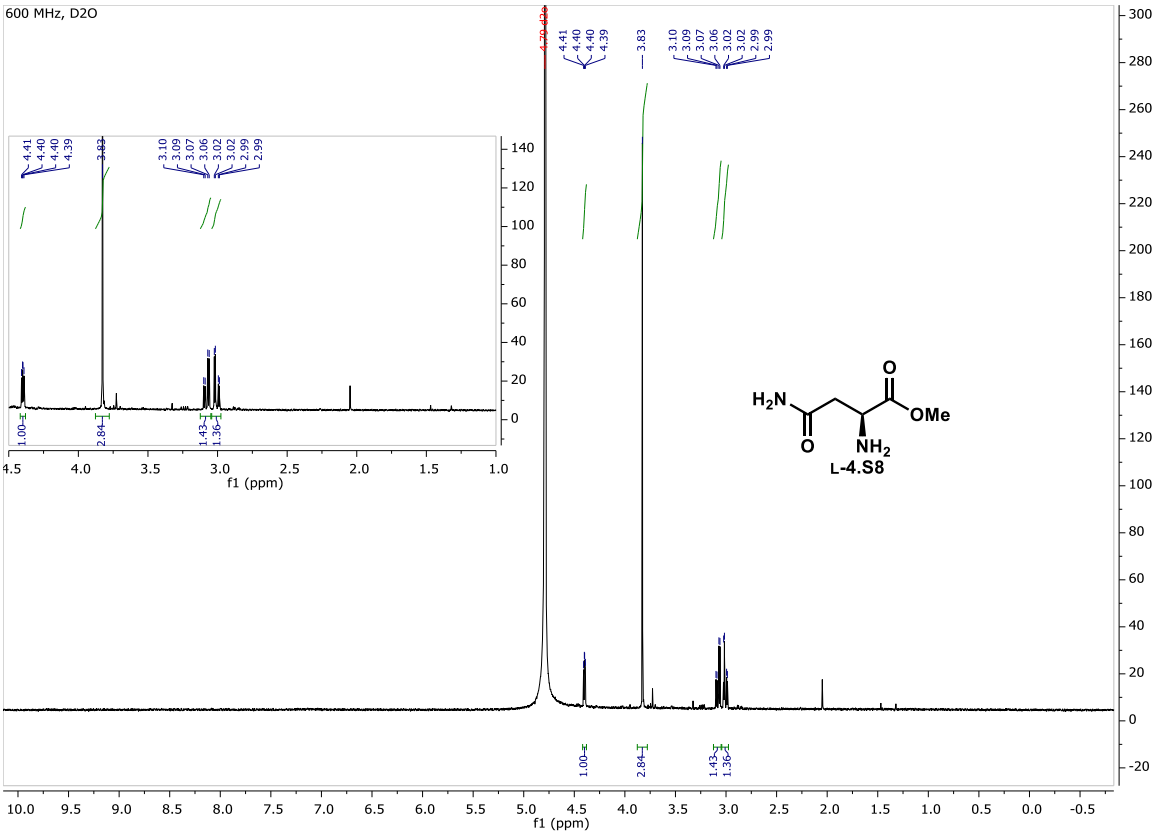

600 MHz, D2O
D-4.S4
2 HCl
f1 (ppm)

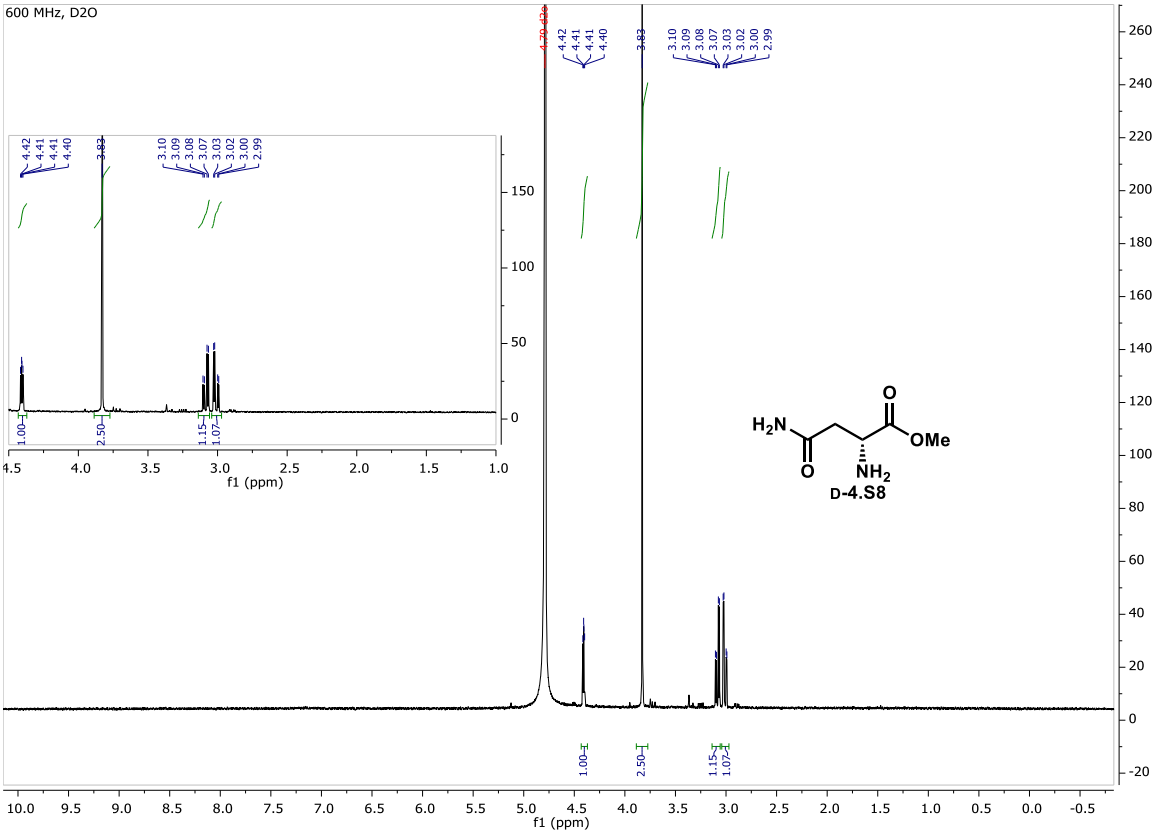

600 MHz, CD3OD
4.87 HDO
D-4.S5
HCl
f1 (ppm)

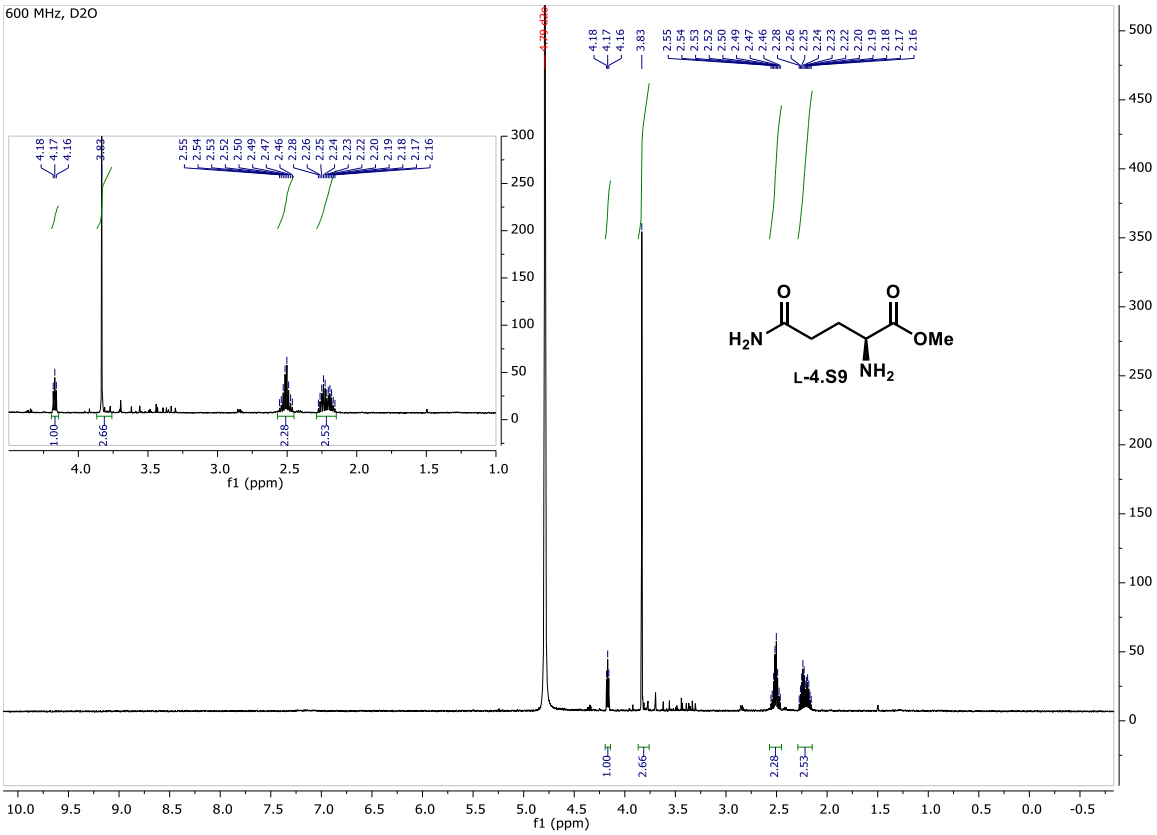

600 MHz, D2O
Me
O
Me
OMe
HCl
NH2
D-4.S6
f1 (ppm)

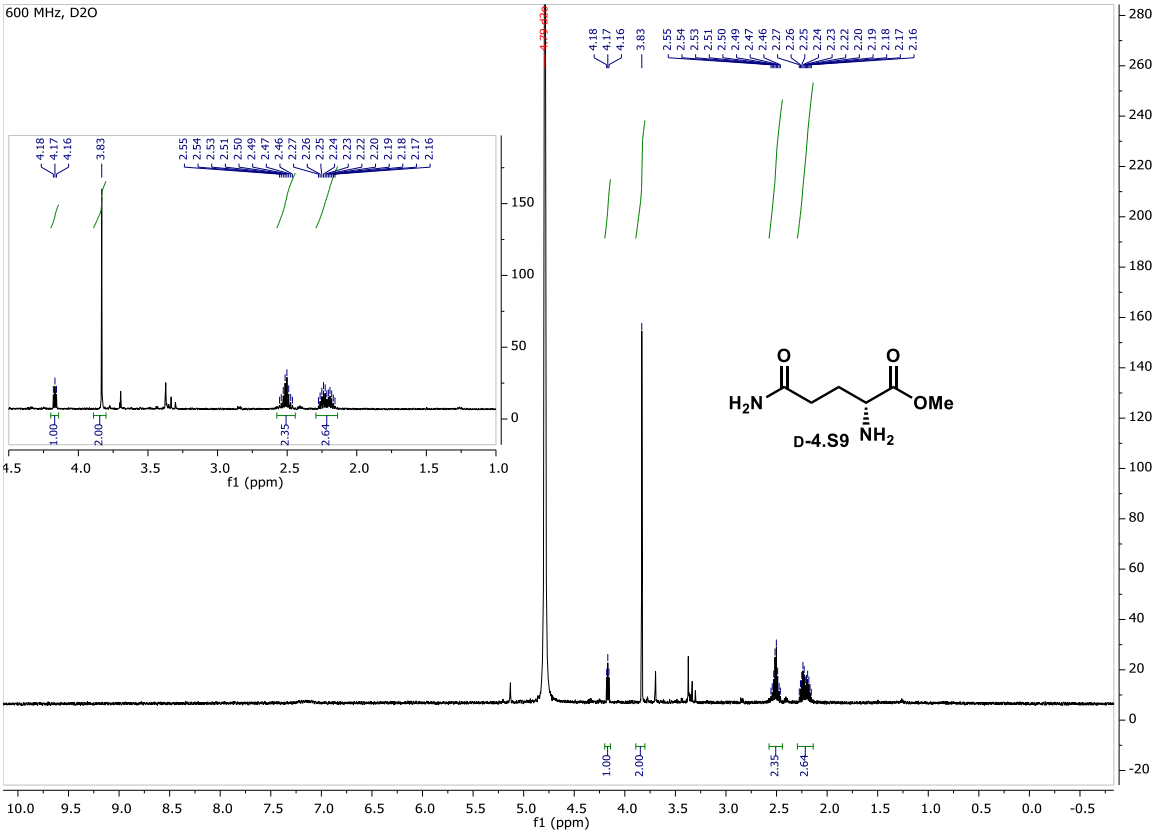

600 MHz, D2O
f1 (ppm)
L-4.S7
HCl

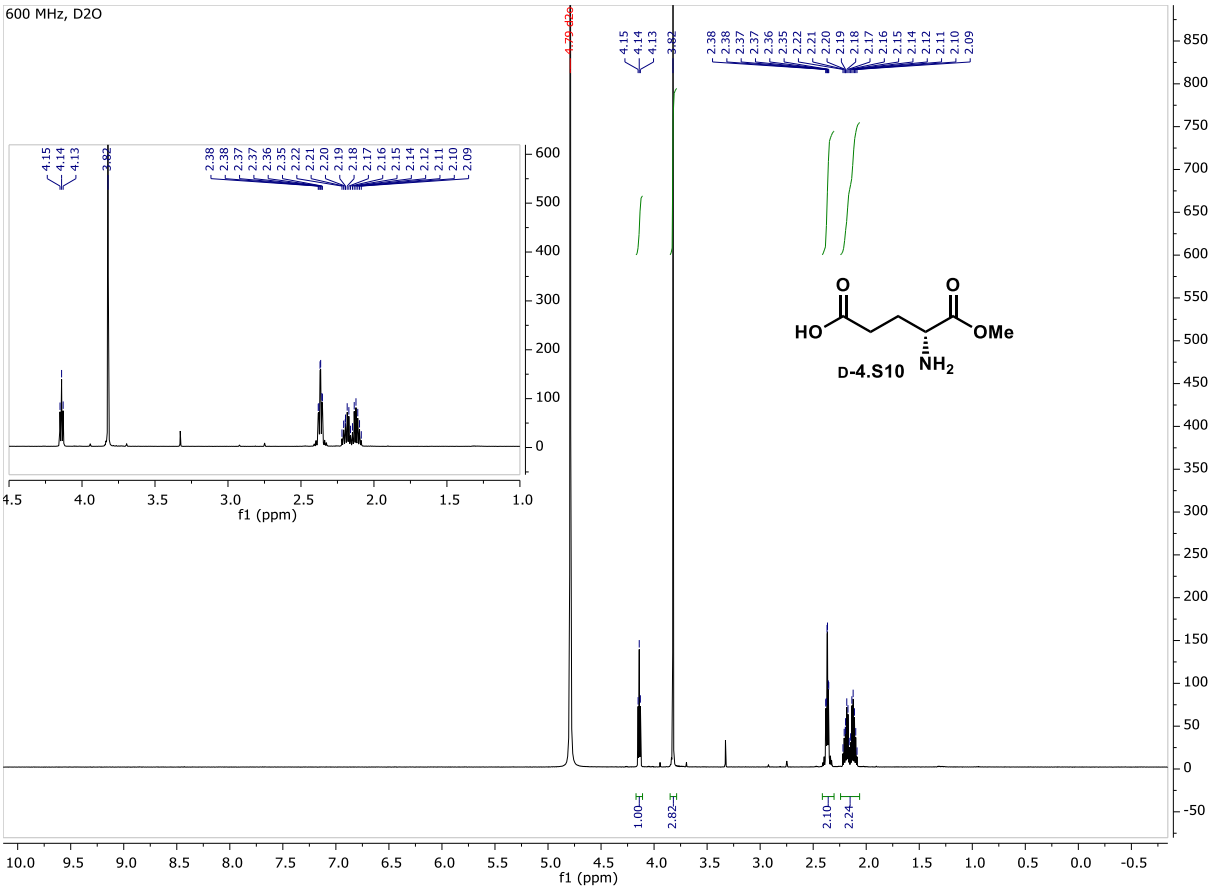

600 MHz, D2O
f1 (ppm)
D-4.S7
HCl

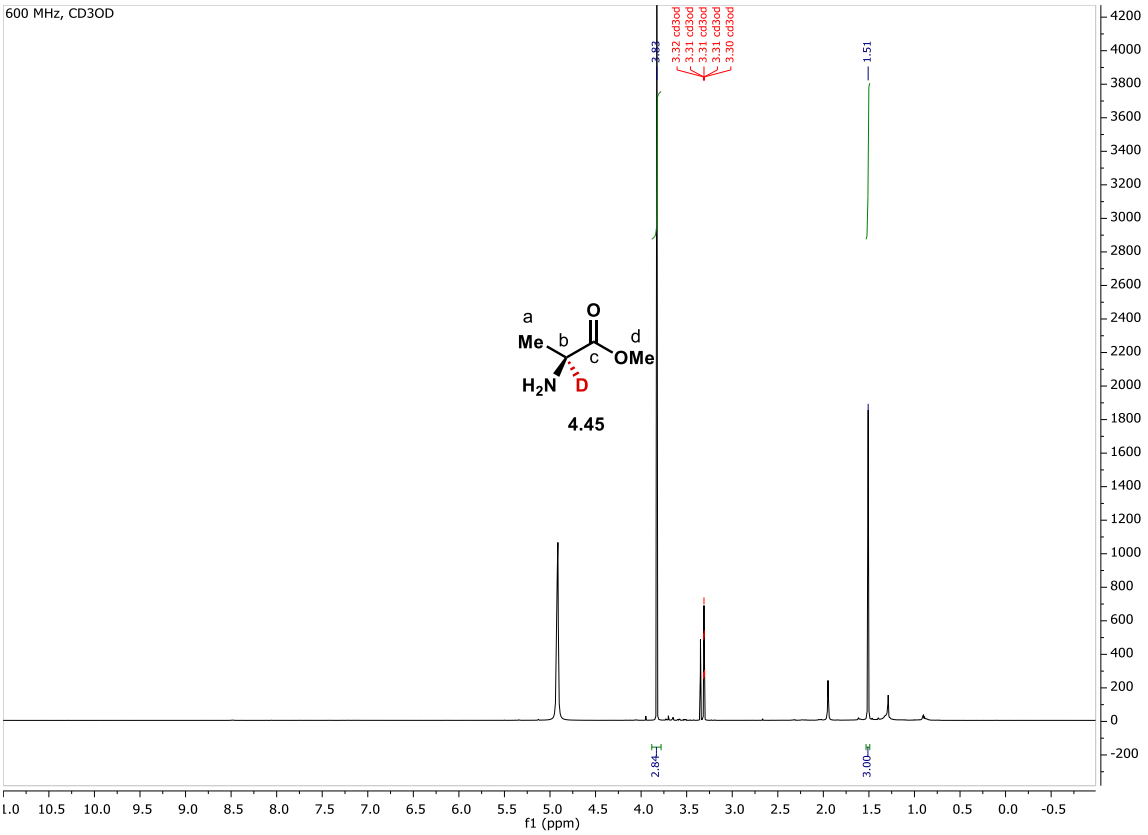

600 MHz, D2O
4.41
4.40
4.40
4.39
3.83
3.10
3.09
3.07
3.06
3.02
3.02
2.99
2.99
1.00
2.84
1.43
1.36
f1 (ppm)
H2N
O
OMe
O
NH2
L-4.S8

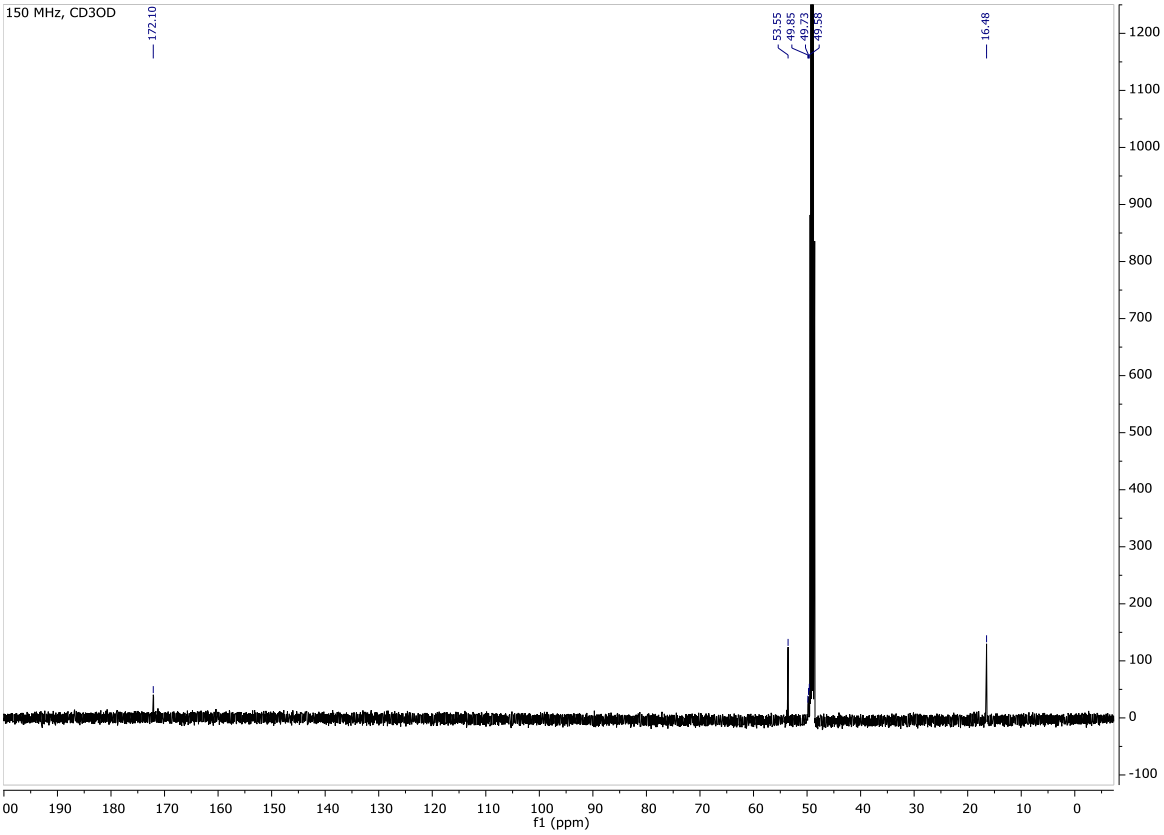

600 MHz, D2O
4.42
4.41
4.41
4.40
3.83
3.10
3.09
3.08
3.07
3.03
3.02
3.00
2.99
1.00
2.50
1.15
1.07
f1 (ppm)
H2N
O
OMe
O
NH2
D-4.S8


600 MHz, D2O
L-4.S9
f1 (ppm)

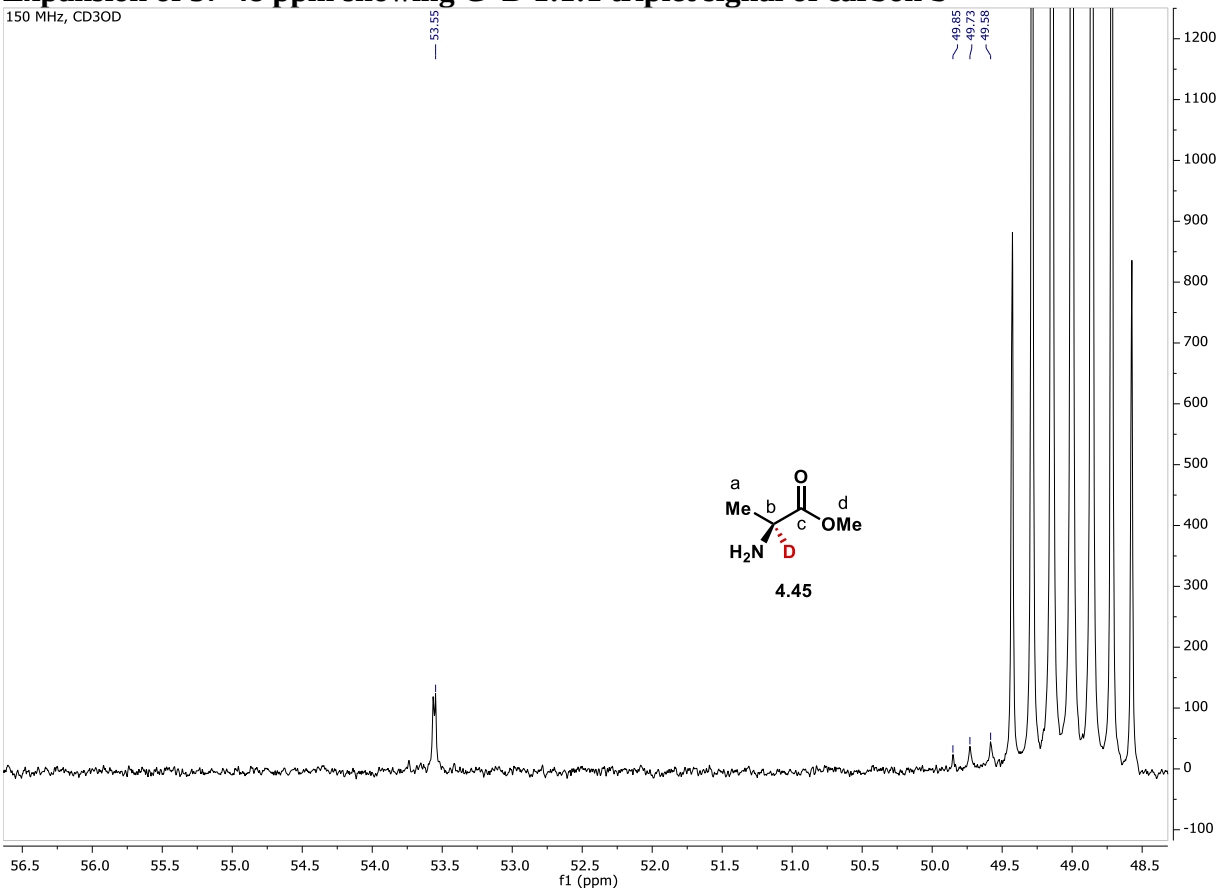

600 MHz, D2O
D-4.S9
f1 (ppm)

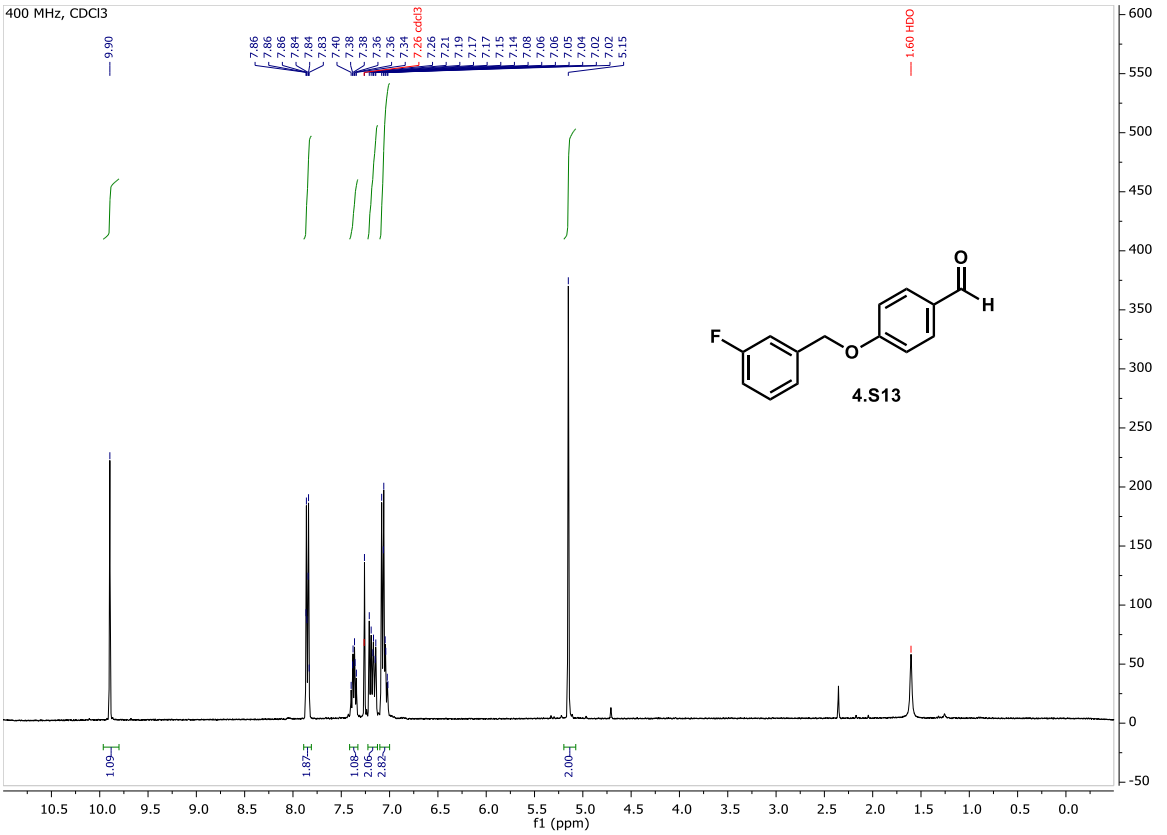

600 MHz, D2O
D-4.S10
f1 (ppm)

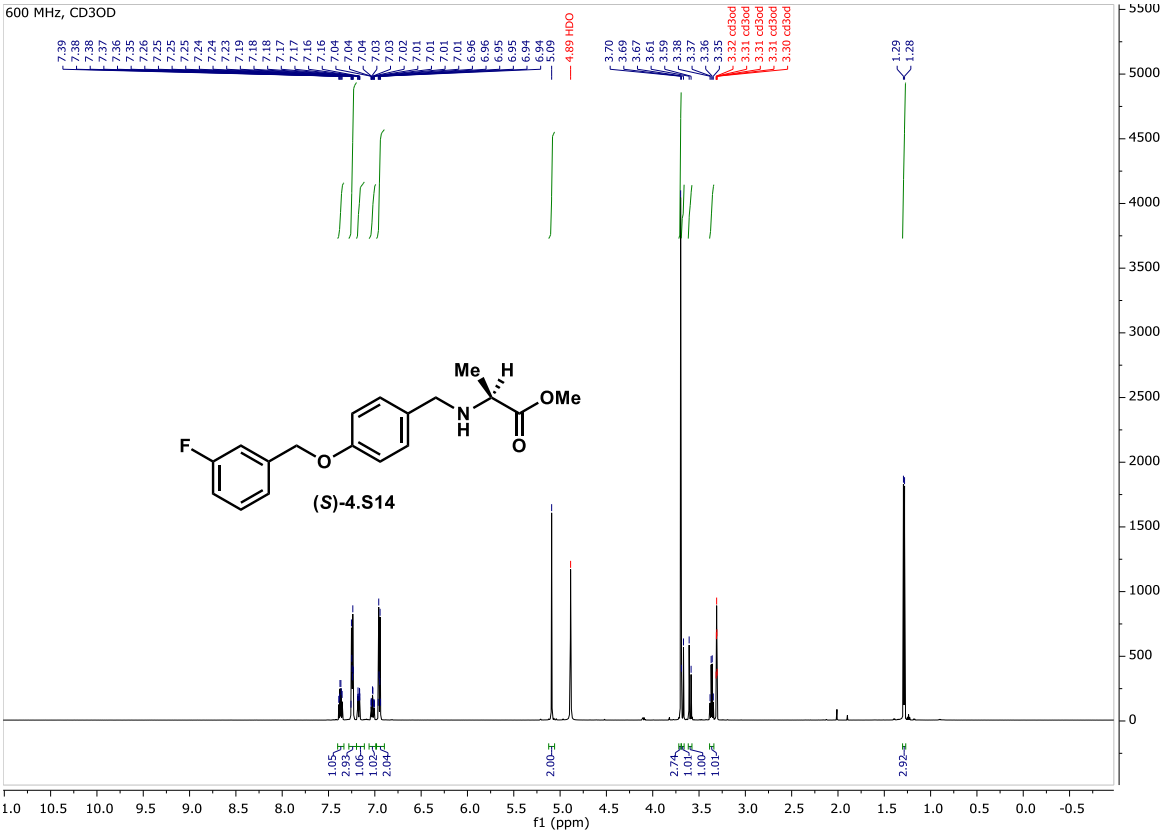

600 MHz, CD3OD
3.32 cd3od
3.31 cd3od
3.31 cd3od
3.31 cd3od
3.30 cd3od
1.51
a
Me
b
c
O
d
OMe
H2N
D
4.45
2.84
3.00
f1 (ppm)

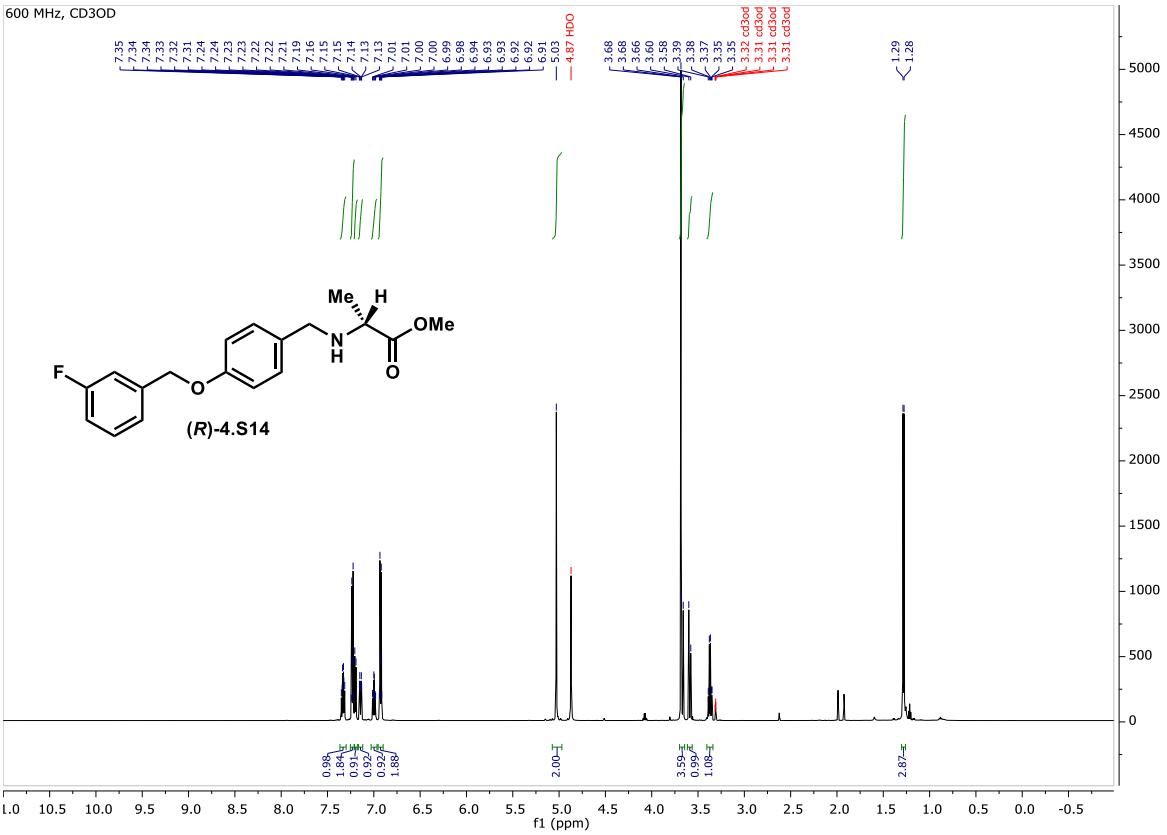

150 MHz, CD3OD
172.10
53.55
49.95
49.73
49.58
16.48
f1 (ppm)

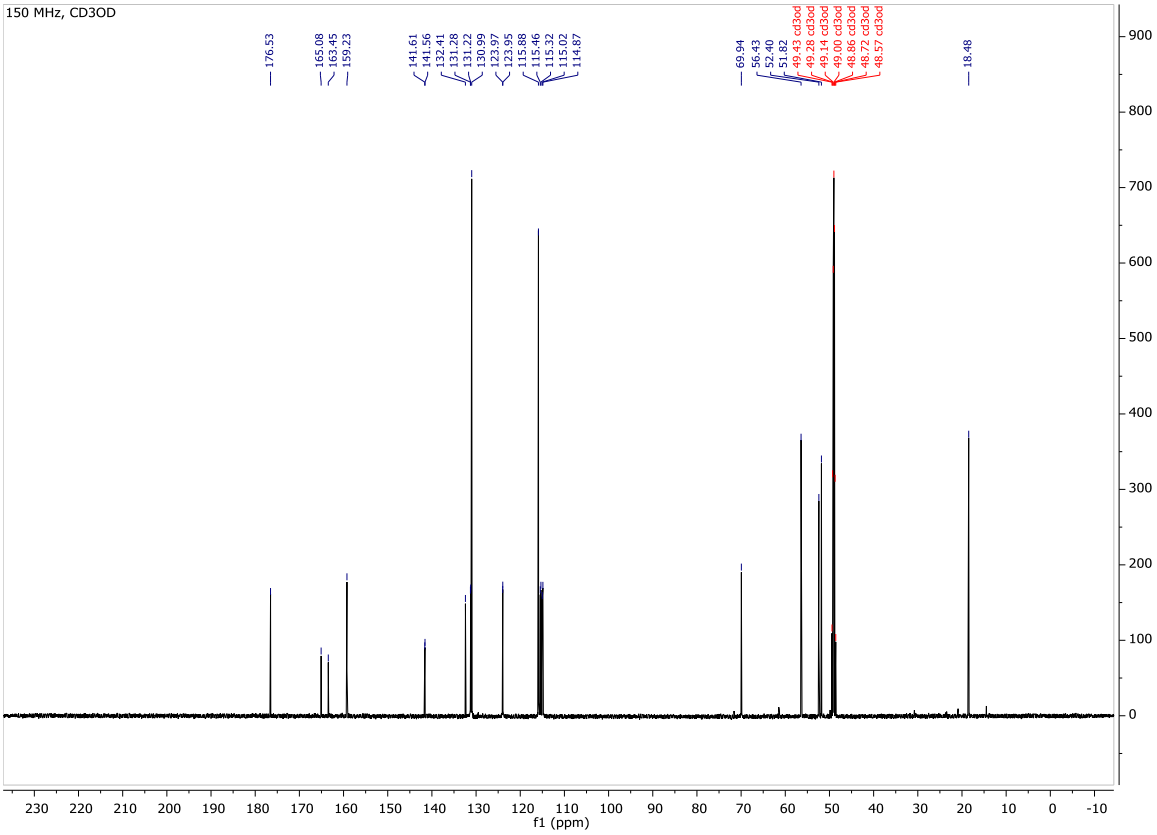
Expansion of 57-48 ppm showing C–D 1:1:1 triplet signal of carbon b

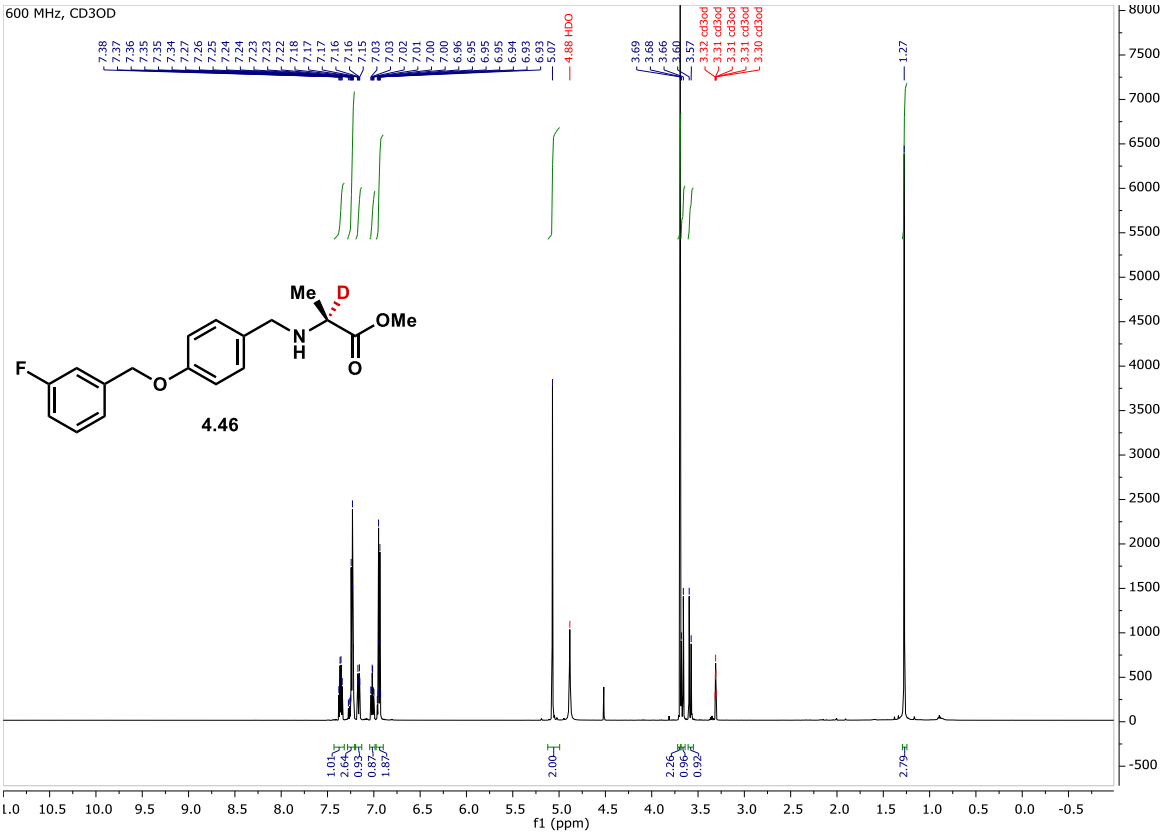


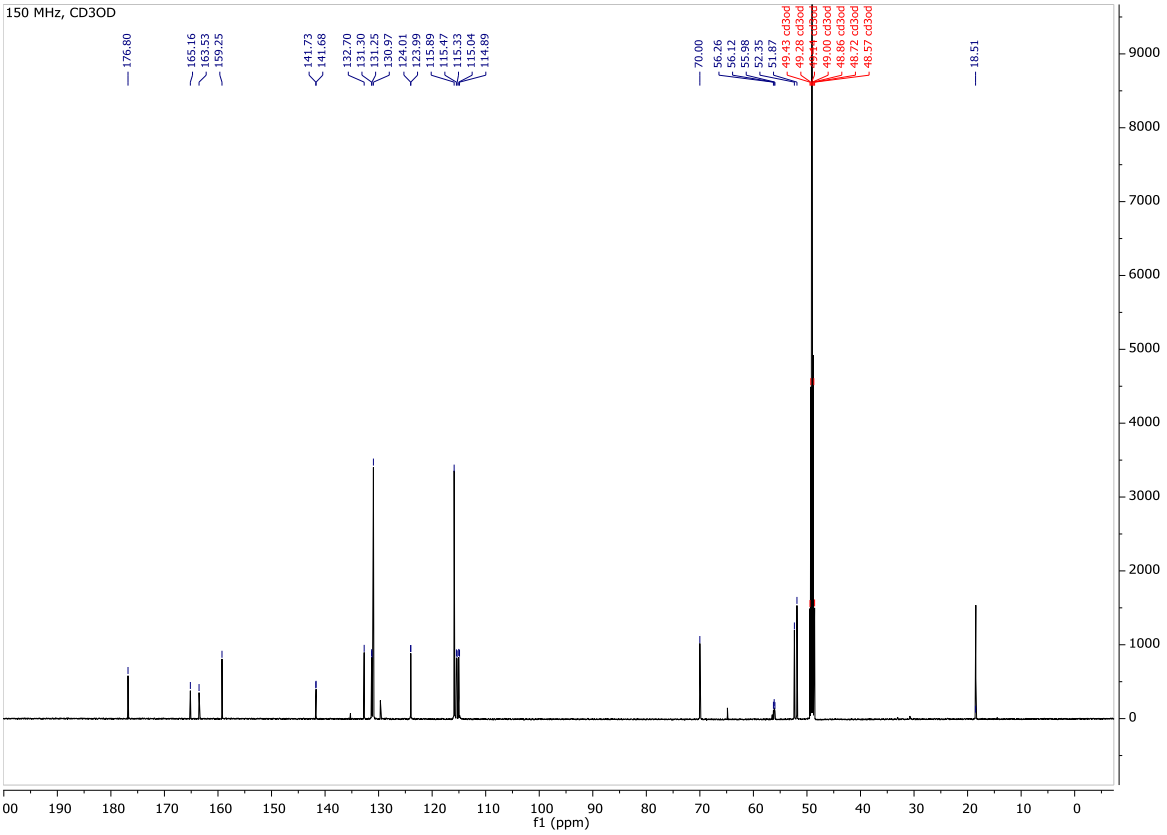

400 MHz, CDCl3
9.90
7.86
7.86
7.86
7.84
7.84
7.83
7.40
7.38
7.38
7.36
7.36
7.34
7.26 cdcl3
7.26
7.21
7.19
7.17
7.17
7.15
7.14
7.08
7.06
7.06
7.05
7.04
7.02
7.02
5.15
1.60 HDO
F
O
O
H
4.S13
1.09
1.87
1.08
2.06
2.82
2.00
f1 (ppm)


600 MHz, CD3OD
7.39
7.38
7.38
7.37
7.36
7.35
7.26
7.25
7.25
7.25
7.24
7.24
7.23
7.19
7.18
7.18
7.17
7.17
7.16
7.16
7.04
7.04
7.04
7.03
7.03
7.02
7.01
7.01
7.01
7.01
6.96
6.96
6.95
6.94
6.94
6.94
5.09
4.89 HDO
3.70
3.69
3.67
3.61
3.59
3.38
3.37
3.36
3.35
3.32 cd3od
3.31 cd3od
3.31 cd3od
3.31 cd3od
3.30 cd3od
1.29
1.28
Me
H
OMe
N
H
O
O
F
(S)-4.S14
1.05
2.93
1.06
1.02
2.04
2.00
2.74
1.01
1.00
1.01
2.92
f1 (ppm)

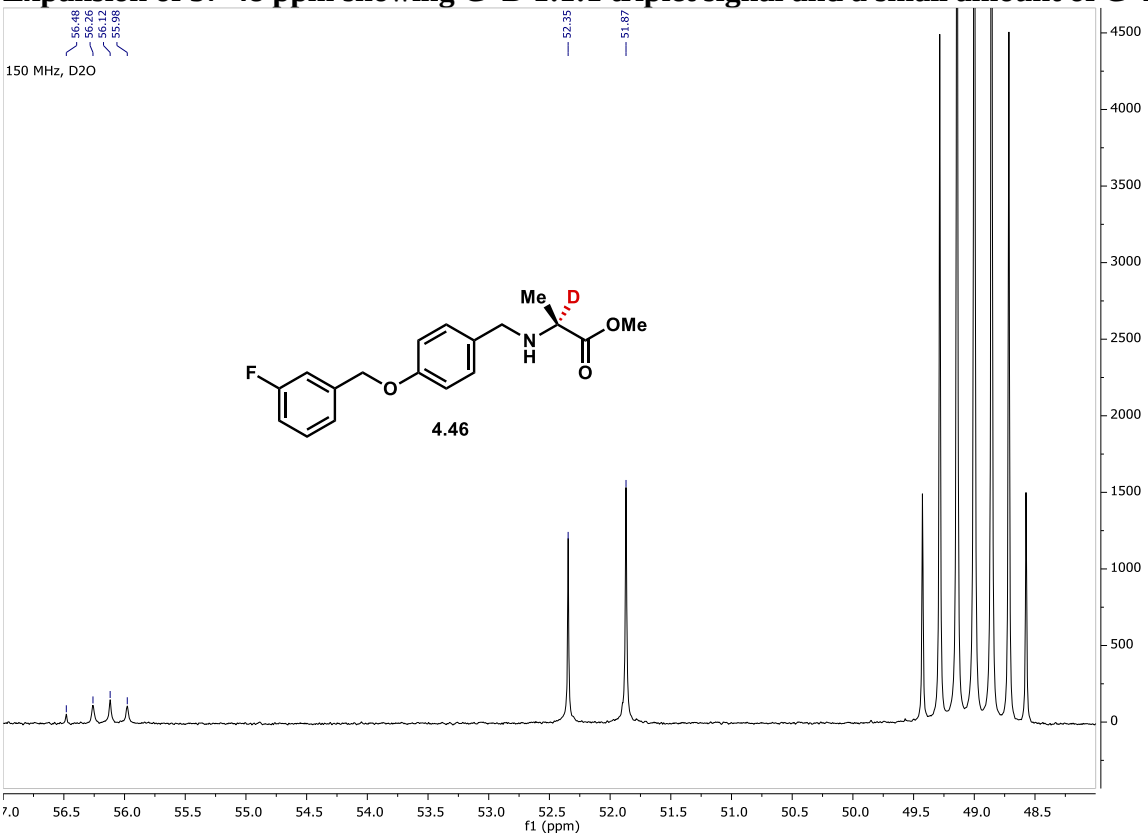

600 MHz, CD3OD
Me
H
N
H
OMe
O
O
F
(R)-4.S14
f1 (ppm)

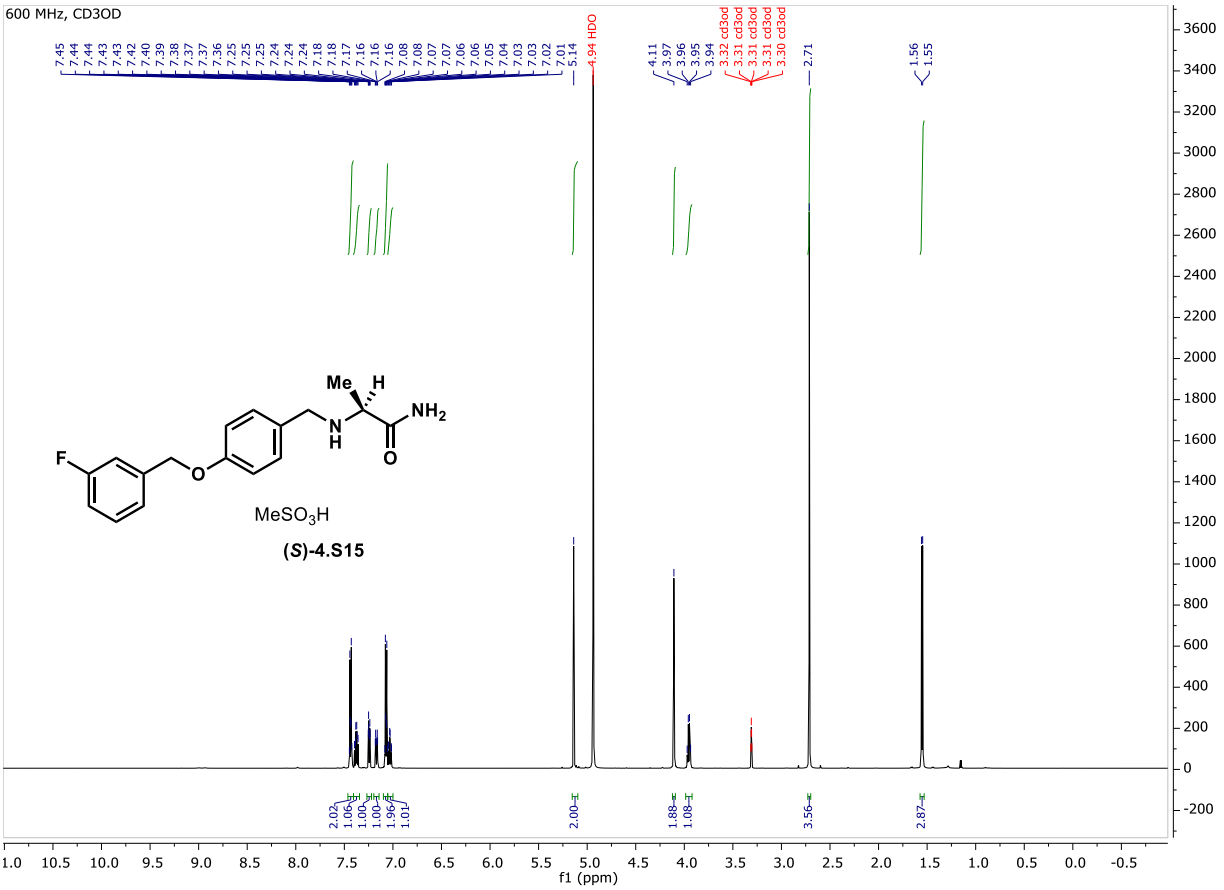

150 MHz, CD3OD
f1 (ppm)

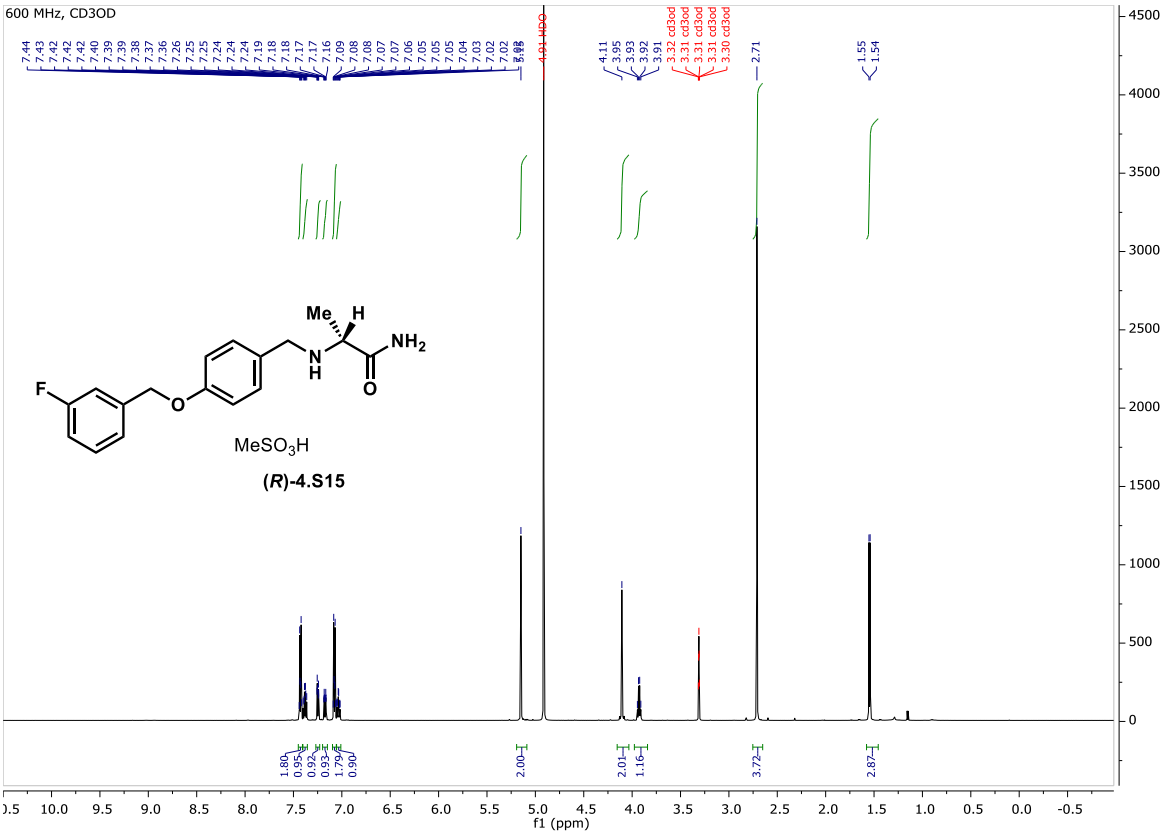

600 MHz, CD3OD
4.88 HDO
Me
D
OMe
N
H
O
F
O
4.46
f1 (ppm)

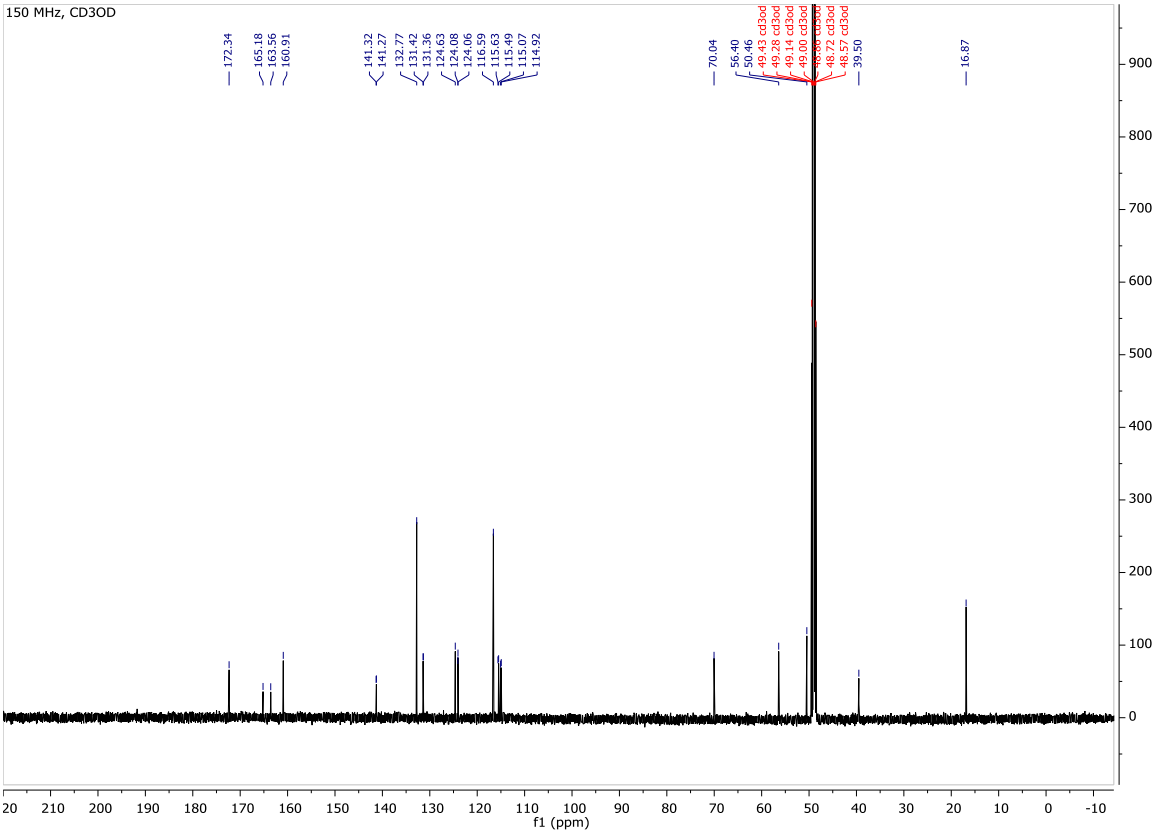

150 MHz, CD3OD
176.80
165.16
163.53
159.25
141.73
141.68
132.70
131.30
131.25
130.97
124.01
123.99
115.89
115.47
115.33
115.04
114.89
70.00
56.26
56.12
55.98
52.35
51.87
18.51
f1 (ppm)

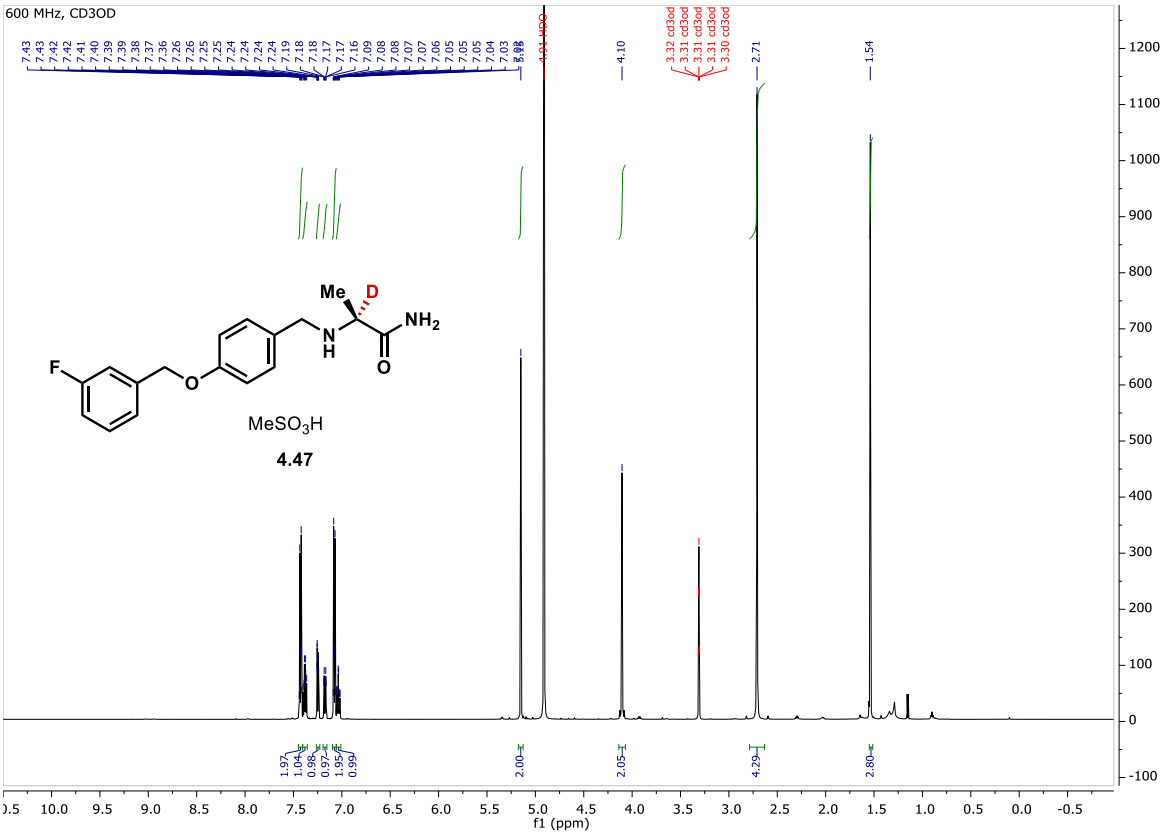
Expansion of 57-48 ppm showing C–D 1:1:1 triplet signal and a small amount of C–H

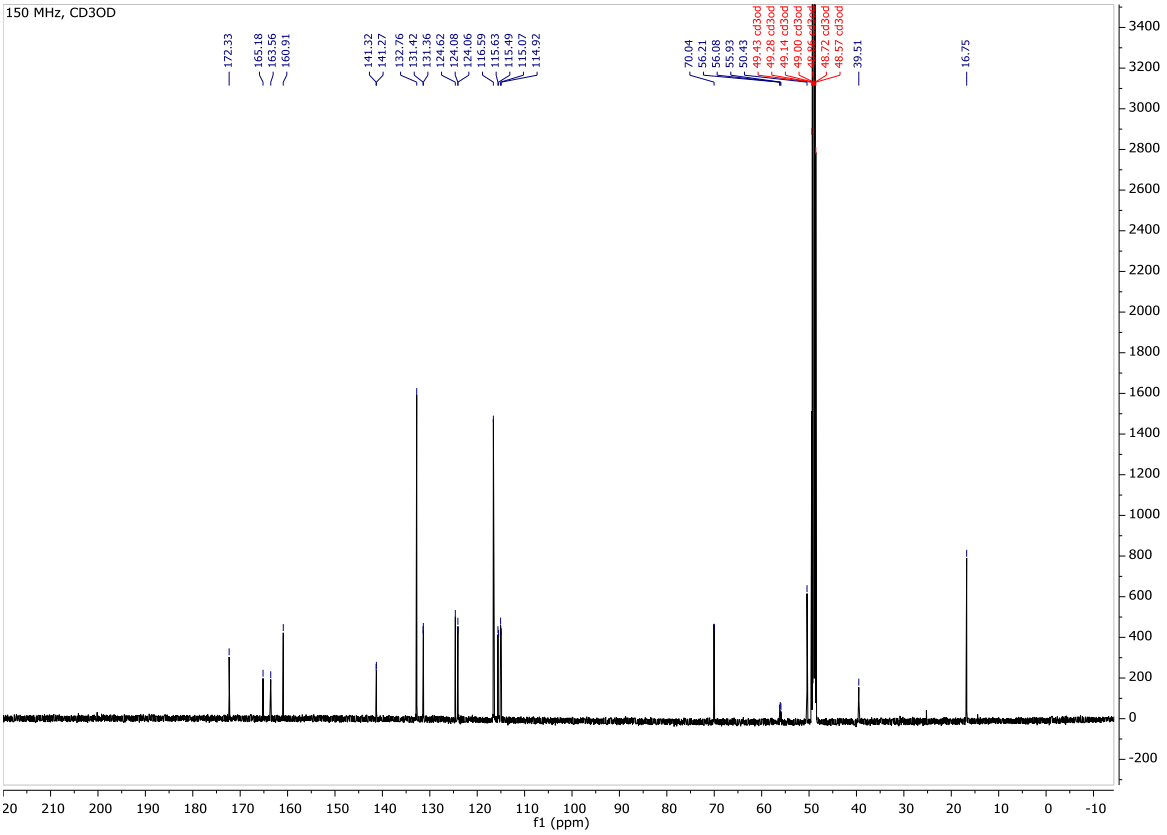



600 MHz, CD3OD
4.94 HDO
3.32 cd3od
3.31 cd3od
3.31 cd3od
3.31 cd3od
3.30 cd3od
2.71
Me
H
NH2
N
H
O
F
O
MeSO3H
(S)-4.S15
f1 (ppm)

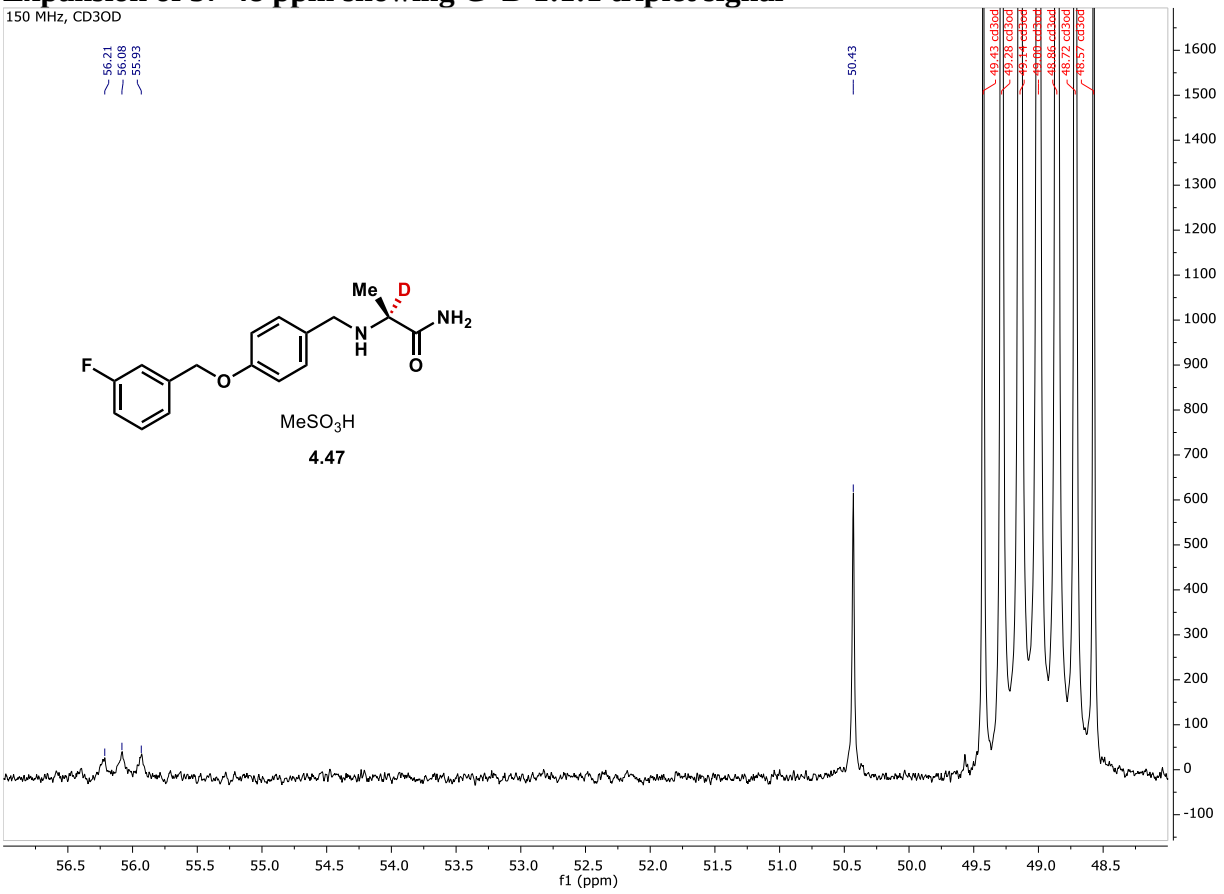

600 MHz, CD3OD
Me
H
NH2
N
H
O
F
O
MeSO3H
(R)-4.S15
f1 (ppm)

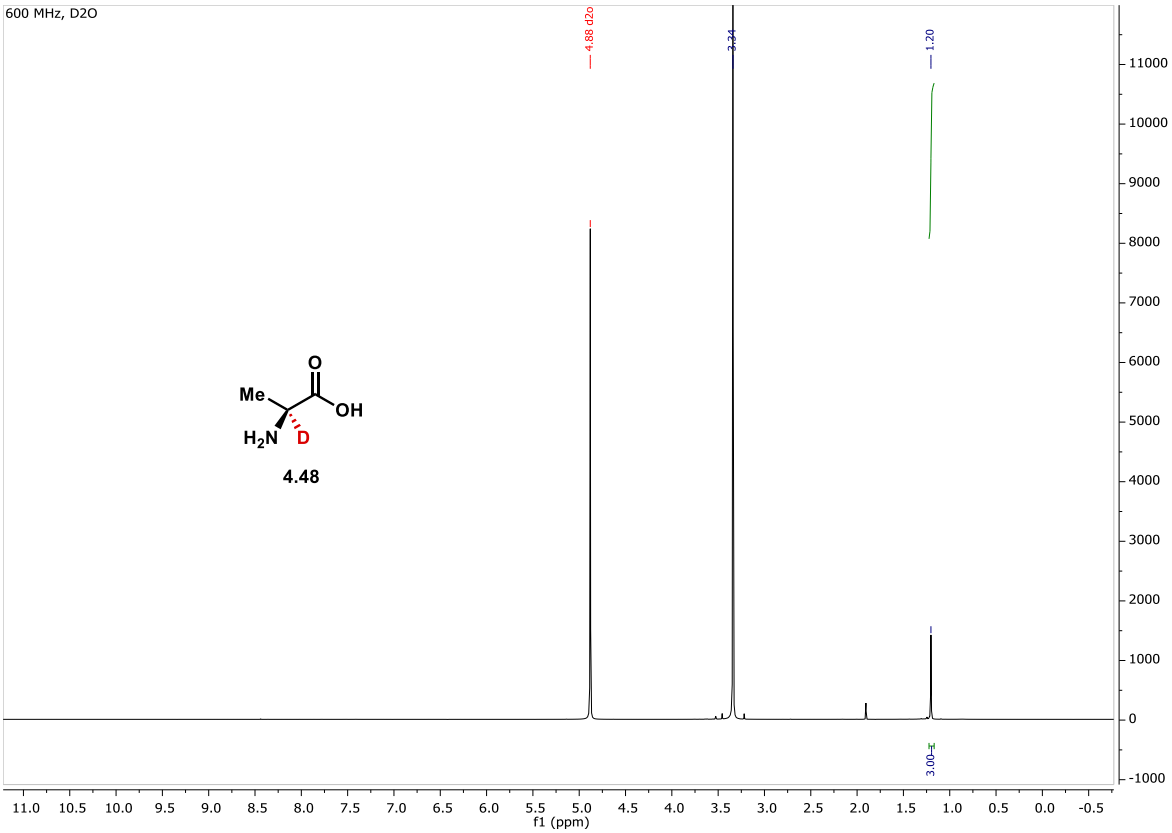

150 MHz, CD3OD
f1 (ppm)

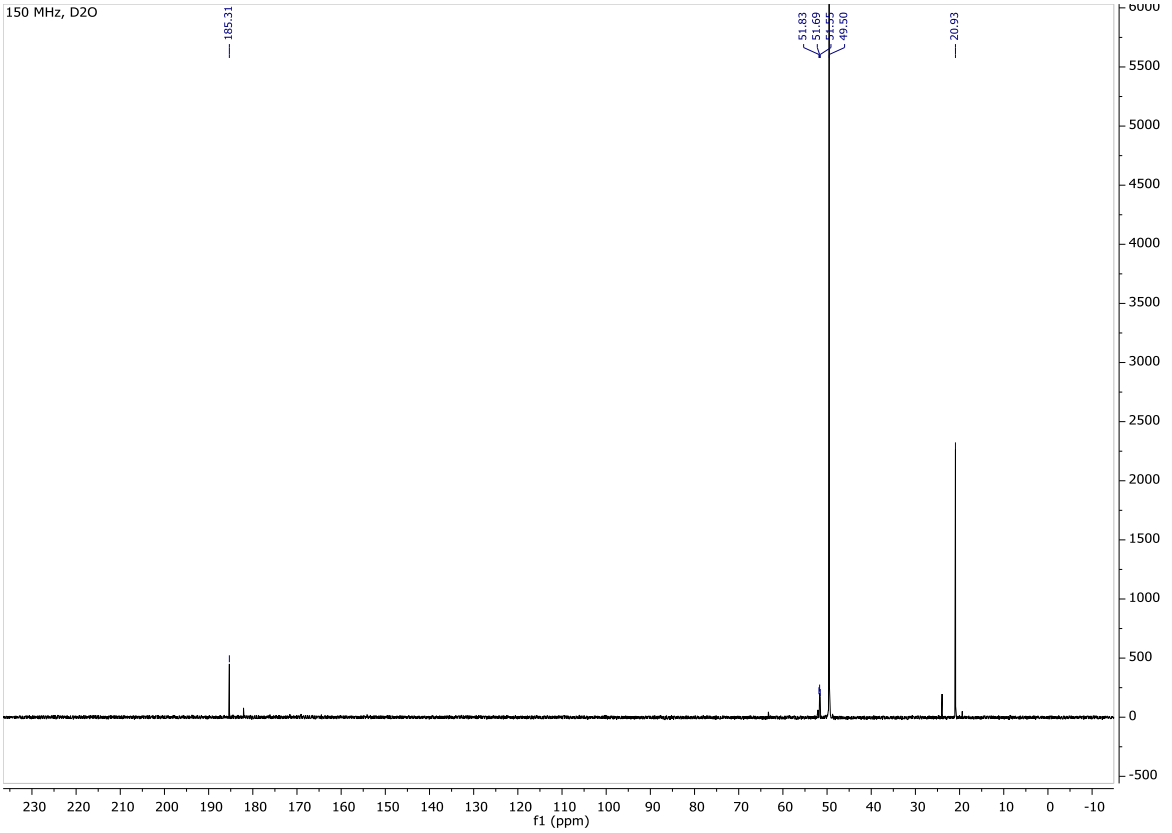

600 MHz, CD3OD
Me
D
NH2
N
H
O
F
O
MeSO3H
4.47
f1 (ppm)


150 MHz, CD3OD
f1 (ppm)

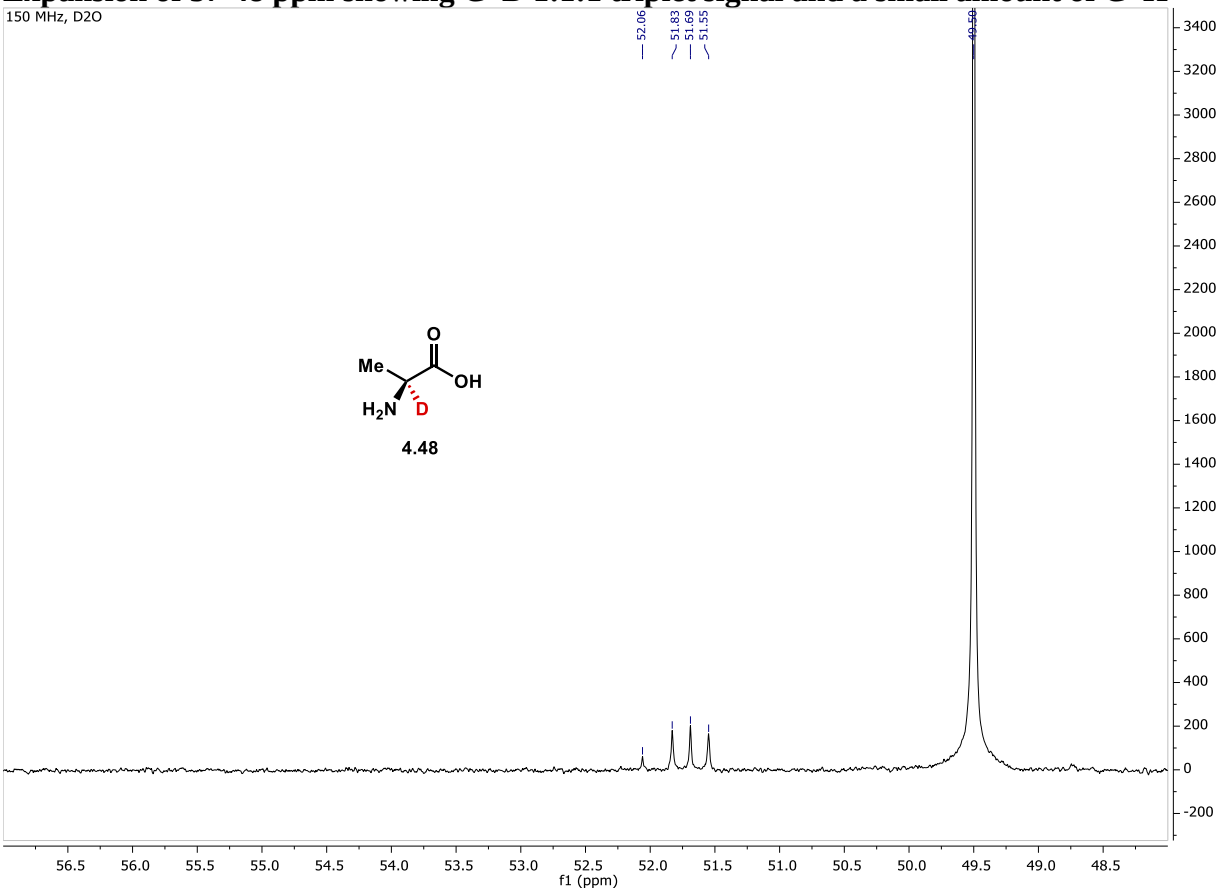
Expansion of 57-48 ppm showing C–D 1:1:1 triplet signal




600 MHz, D2O
4.88 d2o
3.24
1.20
3.00
Me
OH
H2N
D
4.48
f1 (ppm)


150 MHz, D2O
185.31
51.83
51.69
51.55
49.50
20.93
f1 (ppm)

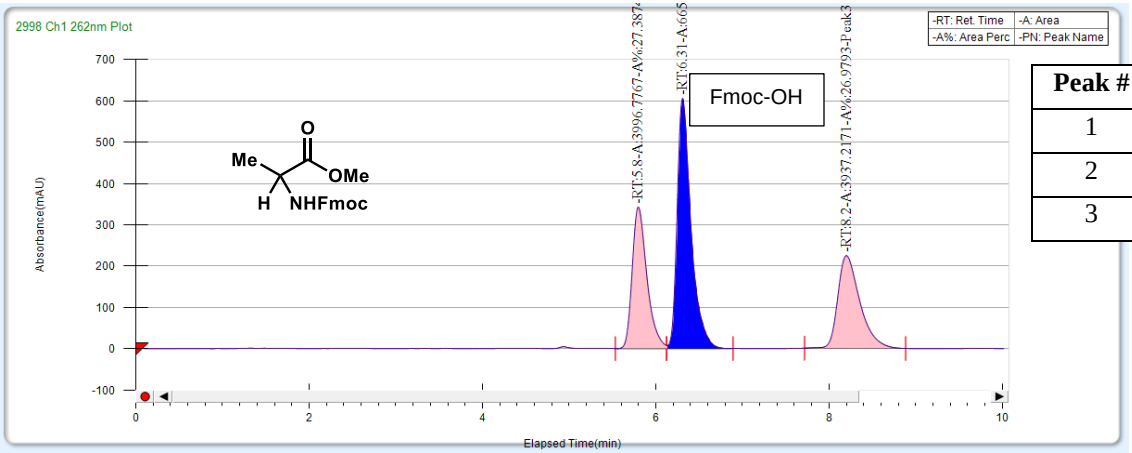
Expansion of 57-48 ppm showing C–D 1:1:1 triplet signal and a small amount of C–H



SFC traces

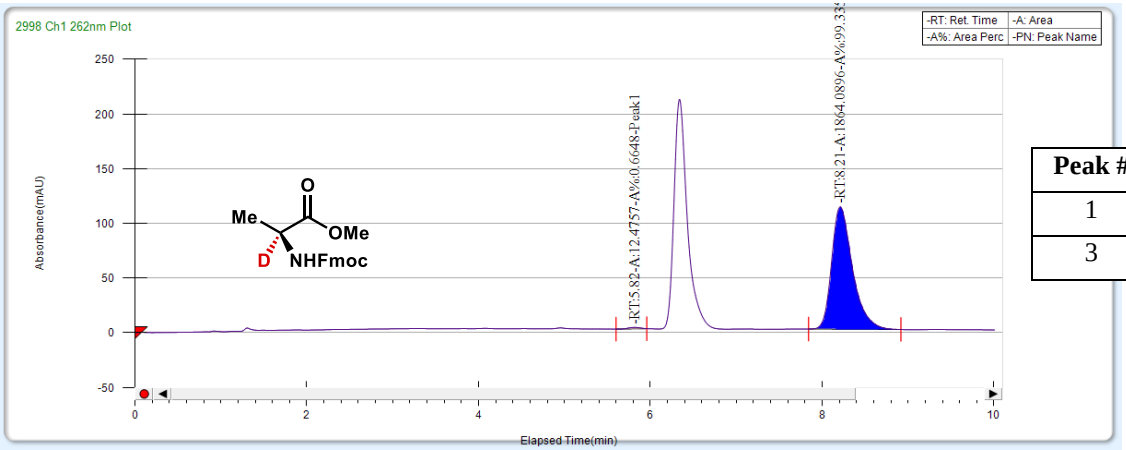
A

Racemic mixture of Ala-OMe



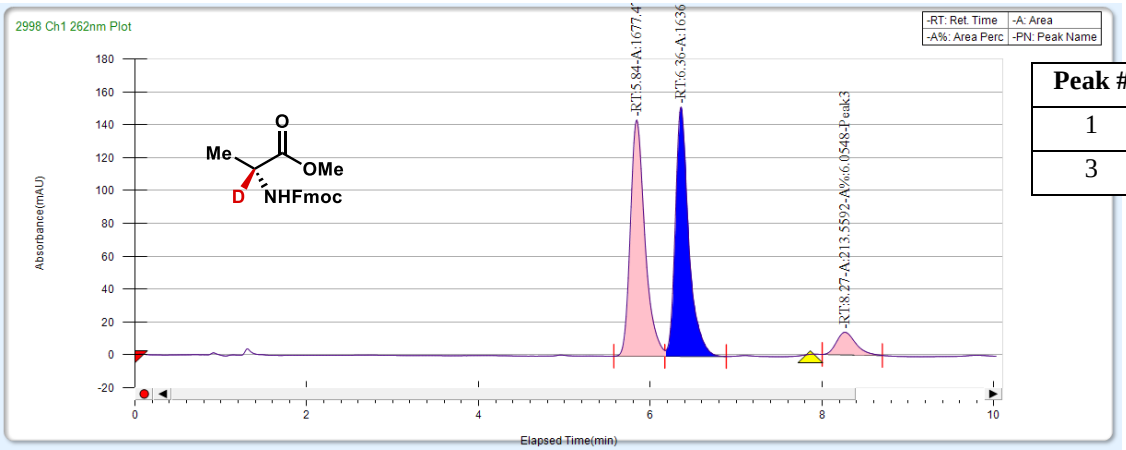
Peak #	Ret. Time	Compound	Area (%)
1	5.8	D-Ala-OMe	27.3874
2	6.3	Fmoc-OH	N/A
3	8.2	L-Ala-OMe	26.9739

L-Ala-OMe incubated with SxtA AONS in D_2O



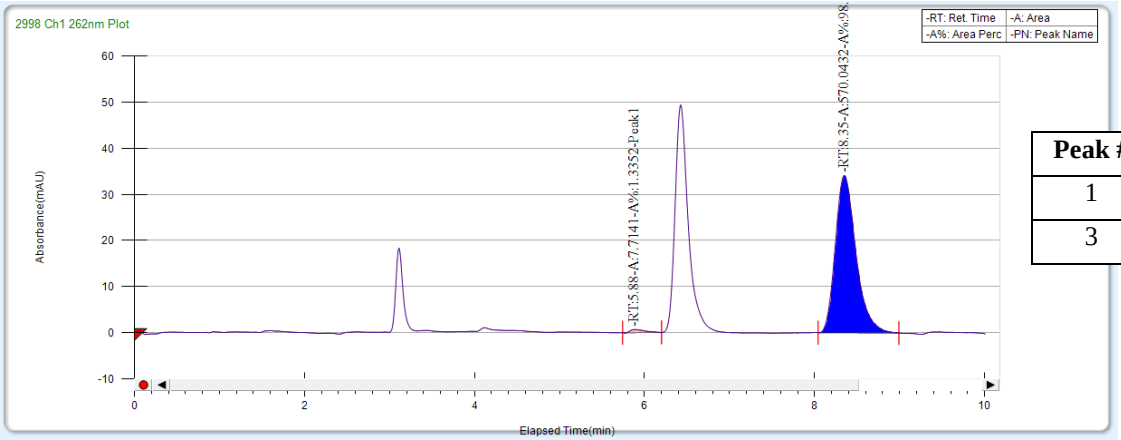
Peak #	Ret. Time	Compound	Area (%)
1	5.8	D-Ala-OMe	0.5916
3	8.2	L-Ala-OMe	99.4084

D-Ala-OMe incubated with SxtA AONS in D_2O



Peak #	Ret. Time	Compound	Area (%)
1	5.8	D-Ala-OMe	47.5593
3	8.2	L-Ala-OMe	6.0548

Preparative-scale L-Ala-OMe incubated with SxtA AONS in D_2O and recovered



Peak #	Ret. Time	Compound	Area (%)
1	5.9	D-Ala-OMe	1.3352
3	8.4	L-Ala-OMe	98.6648

B

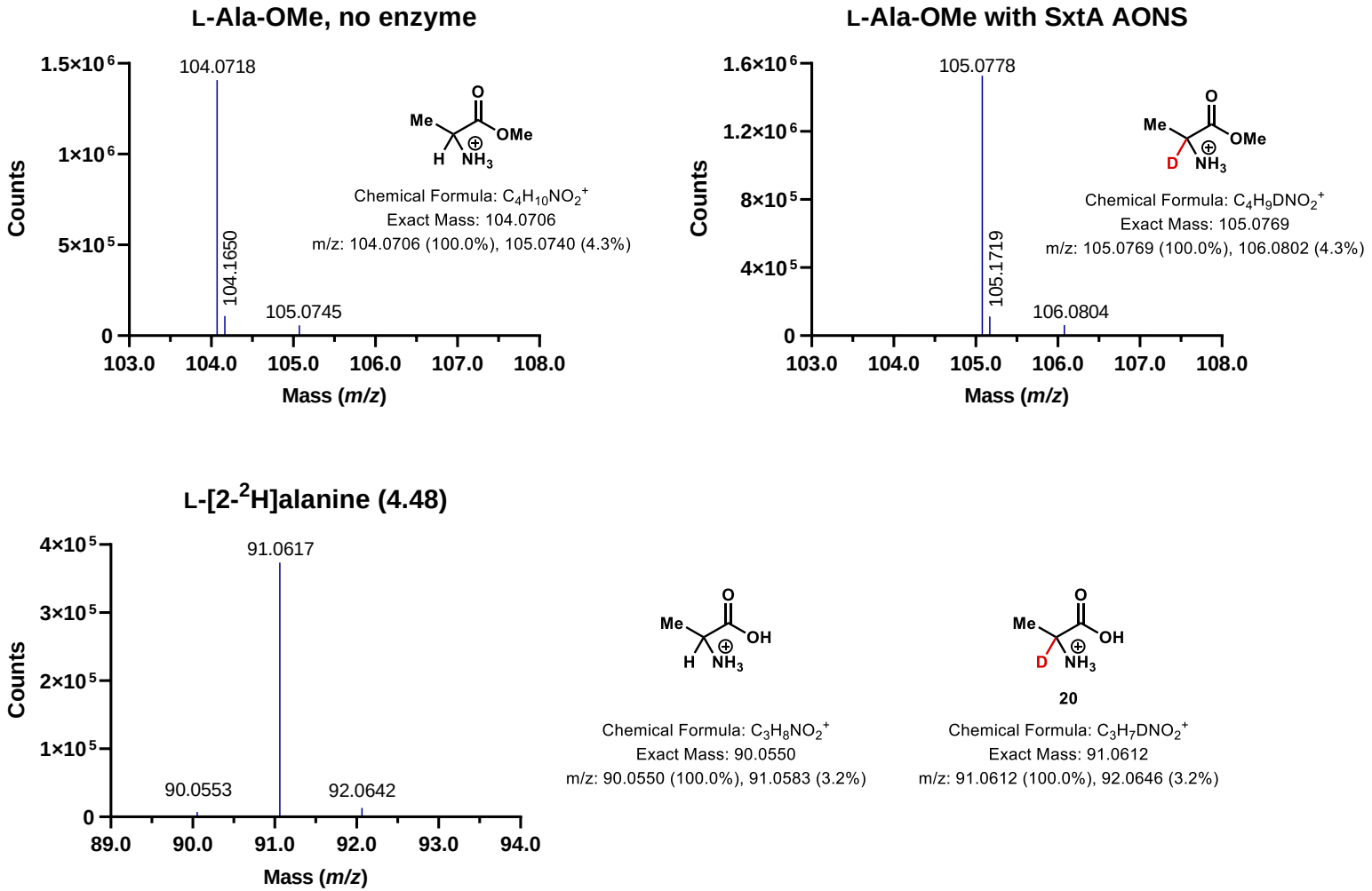
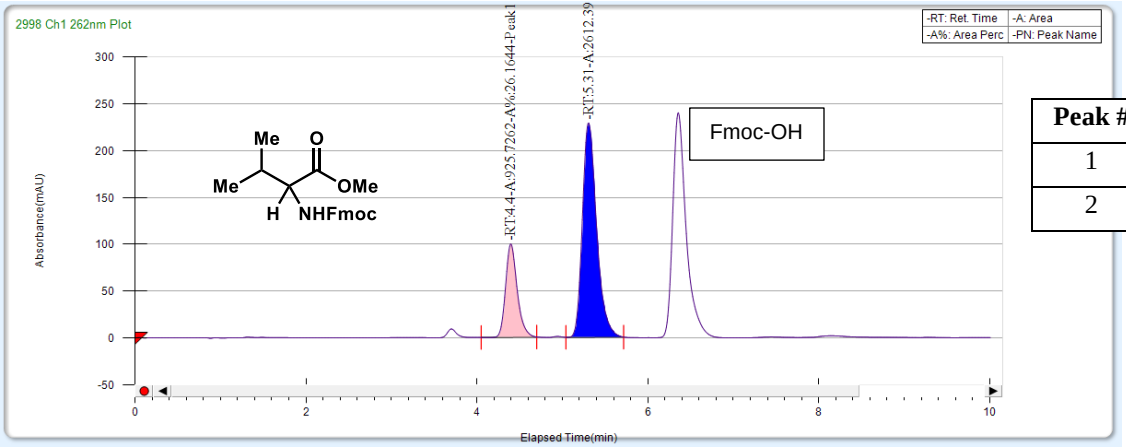


Figure 4.S7. Spectra of alanine and alanine methyl esters.
(A) PDA traces of Fmoc-protected DL-alanine methyl ester from commercial standards, and protected ester after incubation with SxtA AONS (CHIRALCEL OD-H, 15% MeOH with 0.5% formic acid, 3.5 mL/min); (B) representative mass spectra.

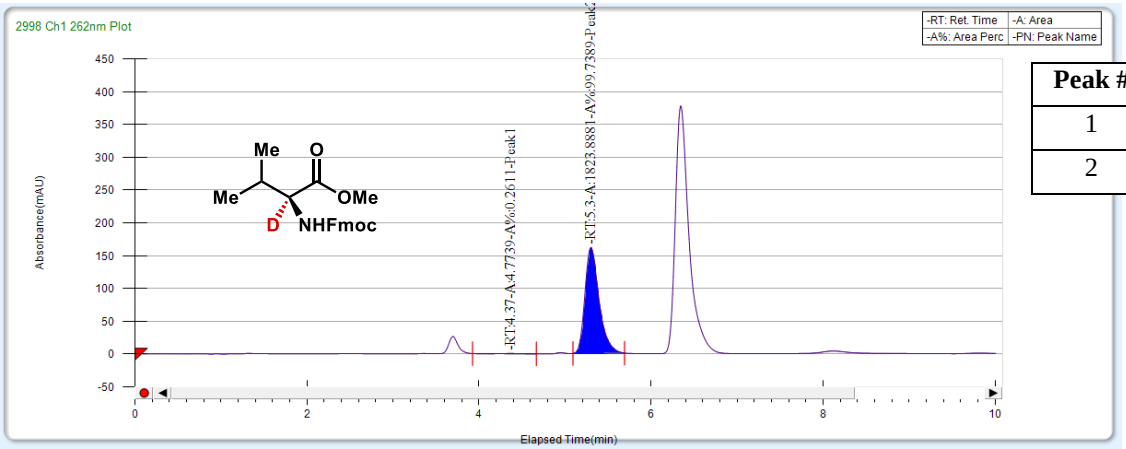
A

Racemic mixture of Val-OMe



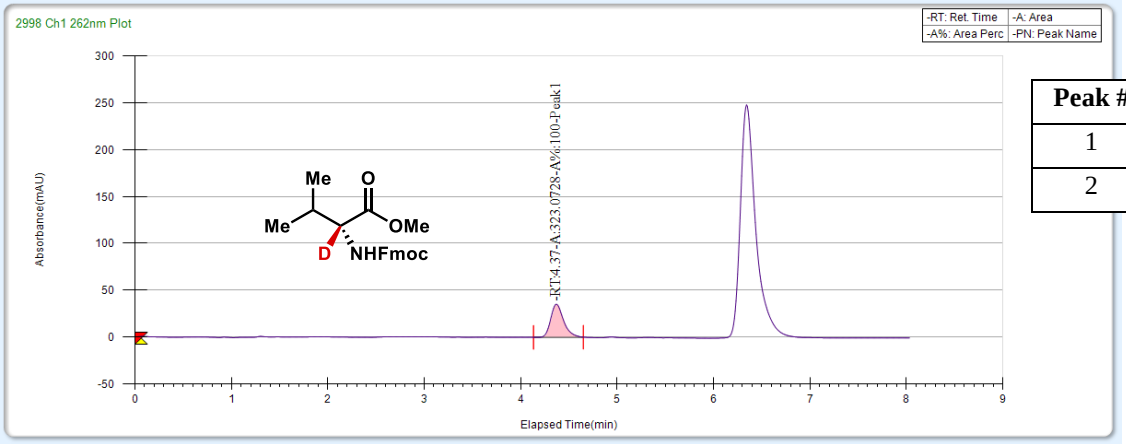
Peak #	Ret. Time	Compound	Area (%)
1	4.4	D-Val-OMe	26.1644
2	5.3	L-Val-OMe	73.8365

L-Val-OMe incubated with SxtA AONS in D_2O



Peak #	Ret. Time	Compound	Area (%)
1	4.4	D-Val-OMe	0.2611
2	5.3	L-Val-OMe	99.7389

D-Val-OMe incubated with SxtA AONS in D_2O



Peak #	Ret. Time	Compound	Area (%)
1	4.4	D-Val-OMe	100
2	5.3	L-Val-OMe	0

B

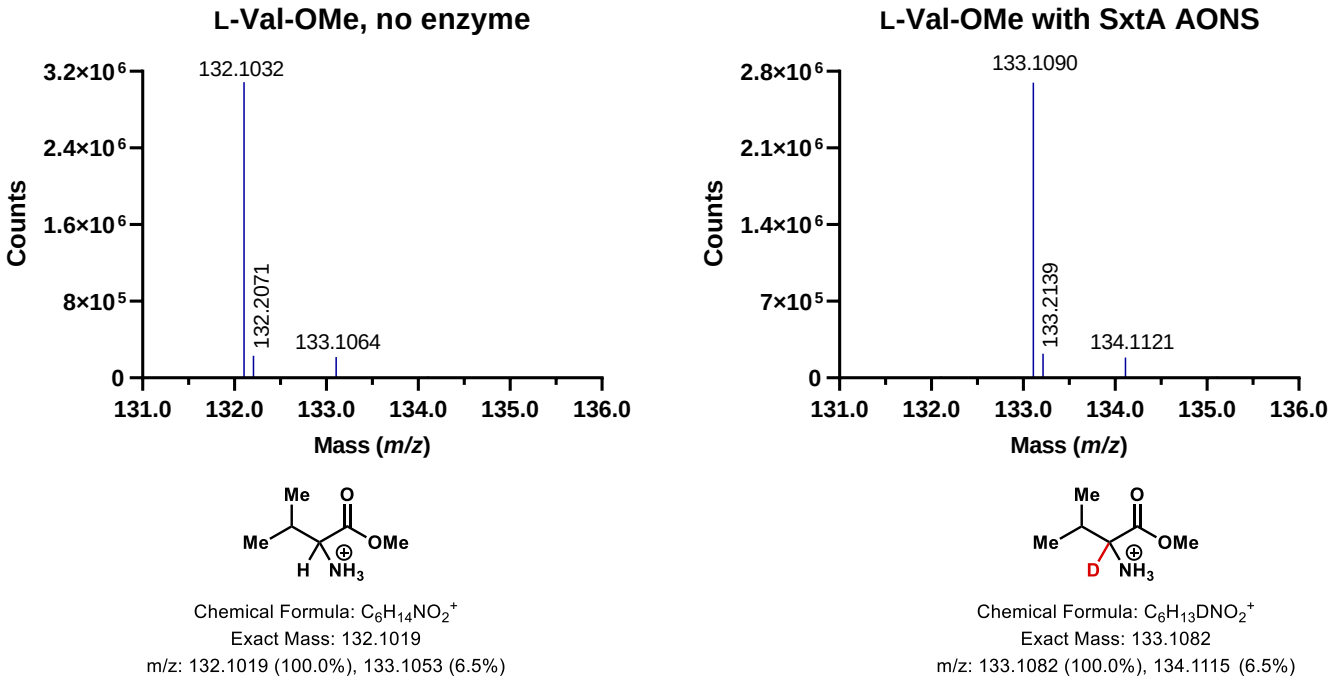
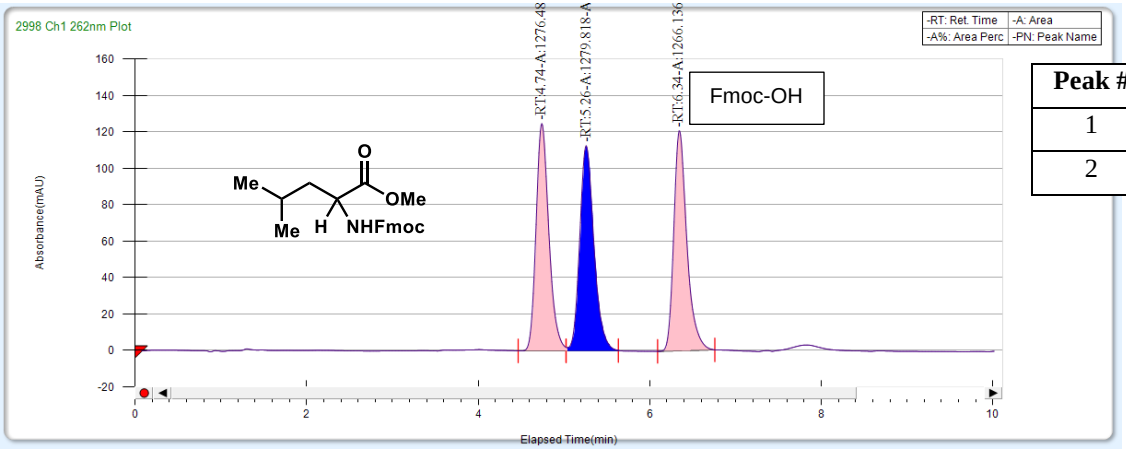


Figure 4.S8. Spectra of valine methyl esters.
(A) PDA traces of Fmoc-protected DL-valine methyl ester from commercial standards, and protected ester after incubation with SxtA AONS (CHIRALCEL OD-H, 15% MeOH with 0.5% formic acid, 3.5 mL/min); (B) representative mass spectra.

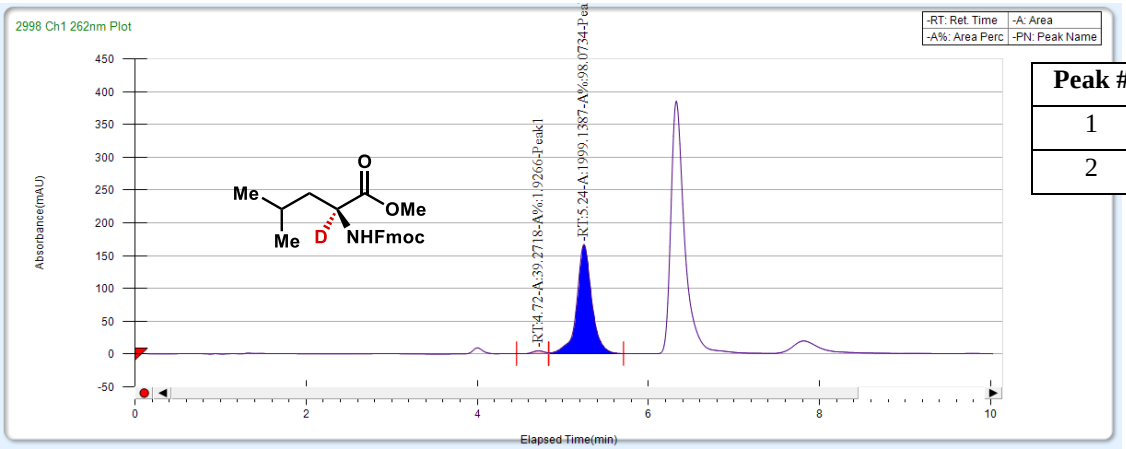
A

Racemic mixture of Leu-OMe



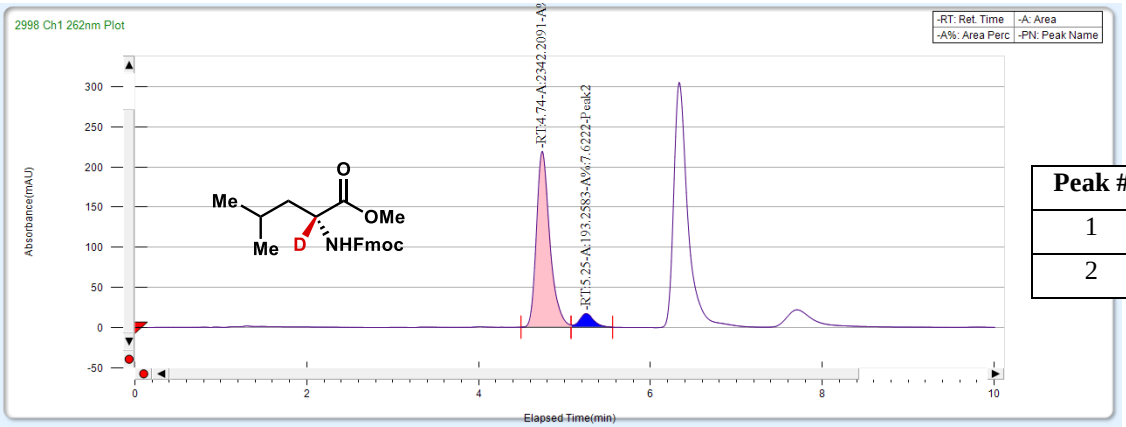
Peak #	Ret. Time	Compound	Area (%)
1	4.7	D-Leu-OMe	33.3945
2	5.3	L-Leu-OMe	33.4817

L-Leu-OMe incubated with SxtA AONS in D_2O



Peak #	Ret. Time	Compound	Area (%)
1	4.7	D-Leu-OMe	1.9266
2	5.3	L-Leu-OMe	98.0734

D-Leu-OMe incubated with SxtA AONS in D_2O



Peak #	Ret. Time	Compound	Area (%)
1	4.7	D-Leu-OMe	92.3778
2	5.3	L-Leu-OMe	7.6222

B

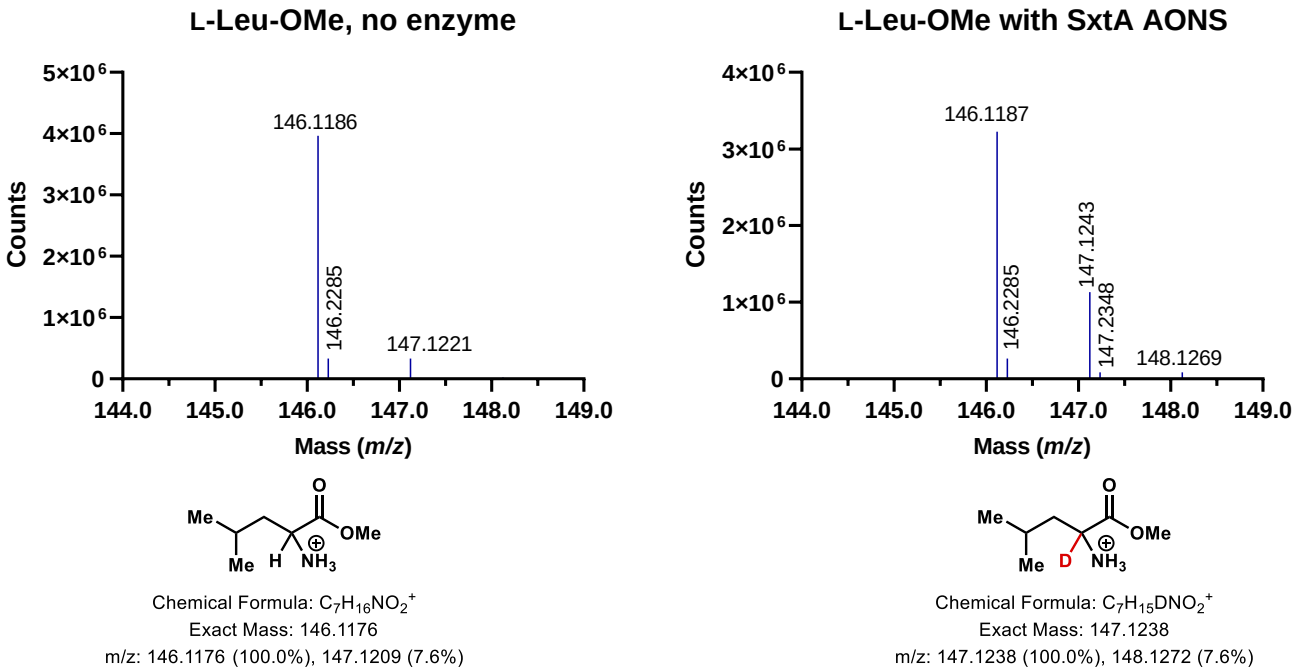
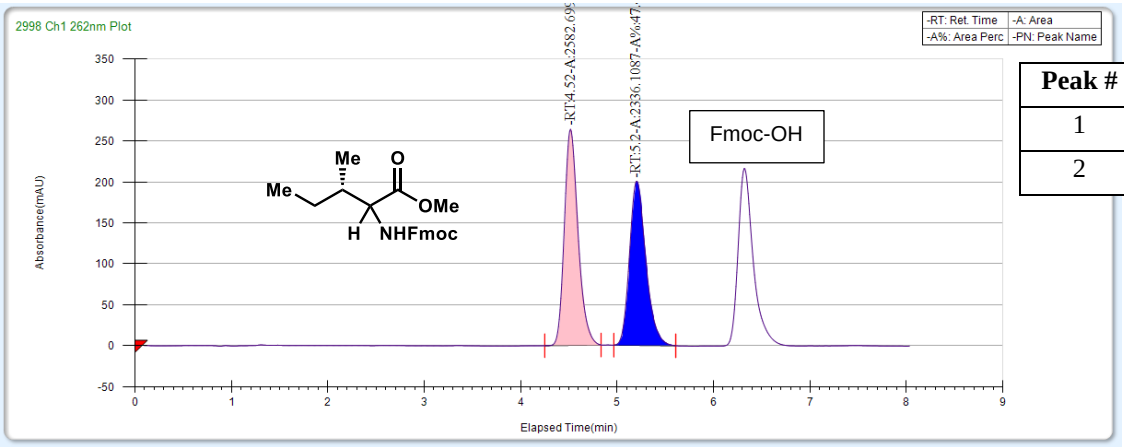


Figure 4.S9. Spectra of leucine methyl esters.
(A) PDA traces of Fmoc-protected DL-leucine methyl ester from commercial standards, and protected ester after incubation with SxtA AONS (CHIRALCEL OD-H, 15% MeOH with 0.5% formic acid, 3.5 mL/min); (B) representative mass spectra.

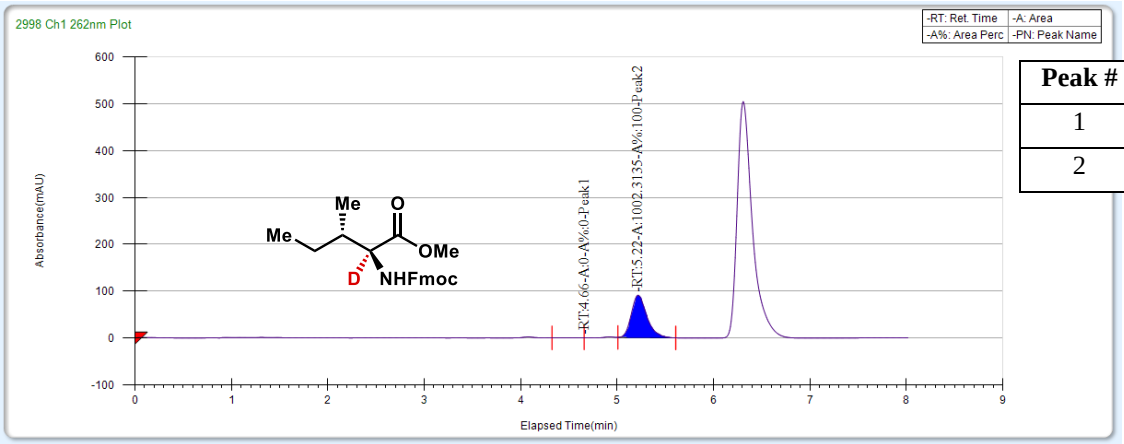
A

Diastereomeric mixture of L-Ile-OMe and D-*allo*-Ile-OMe



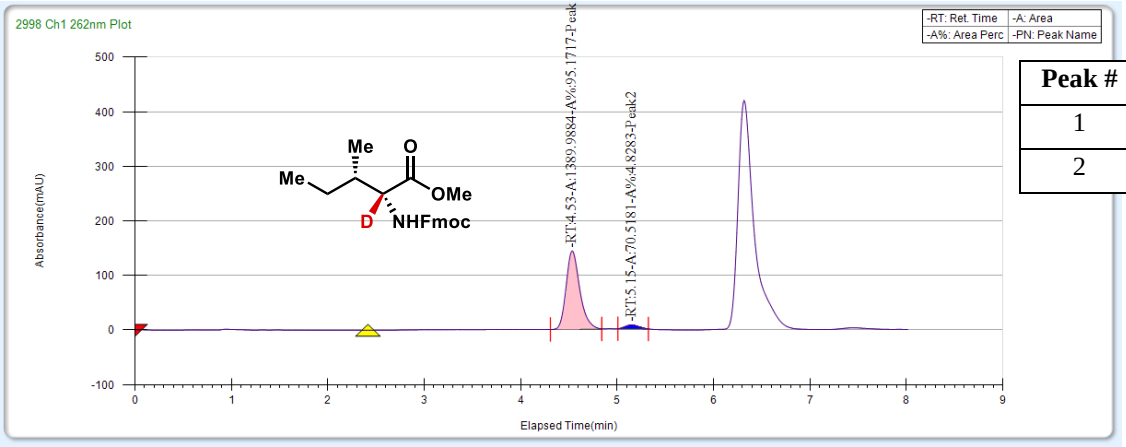
Peak #	Ret. Time	Compound	Area (%)
1	4.5	D-*allo*-Ile-OMe	52.5066
2	5.2	L-Ile-OMe	47.4934

L-Ile-OMe incubated with SxtA AONS in D_2O



Peak #	Ret. Time	Compound	Area (%)
1	4.5	D-*allo*-Ile-OMe	N/A
2	5.2	L-Ile-OMe	100

D-*allo*-Ile-OMe incubated with SxtA AONS in D_2O



Peak #	Ret. Time	Compound	Area (%)
1	4.5	D-*allo*-Ile-OMe	95.1717
2	5.15	contaminant	4.8283

B

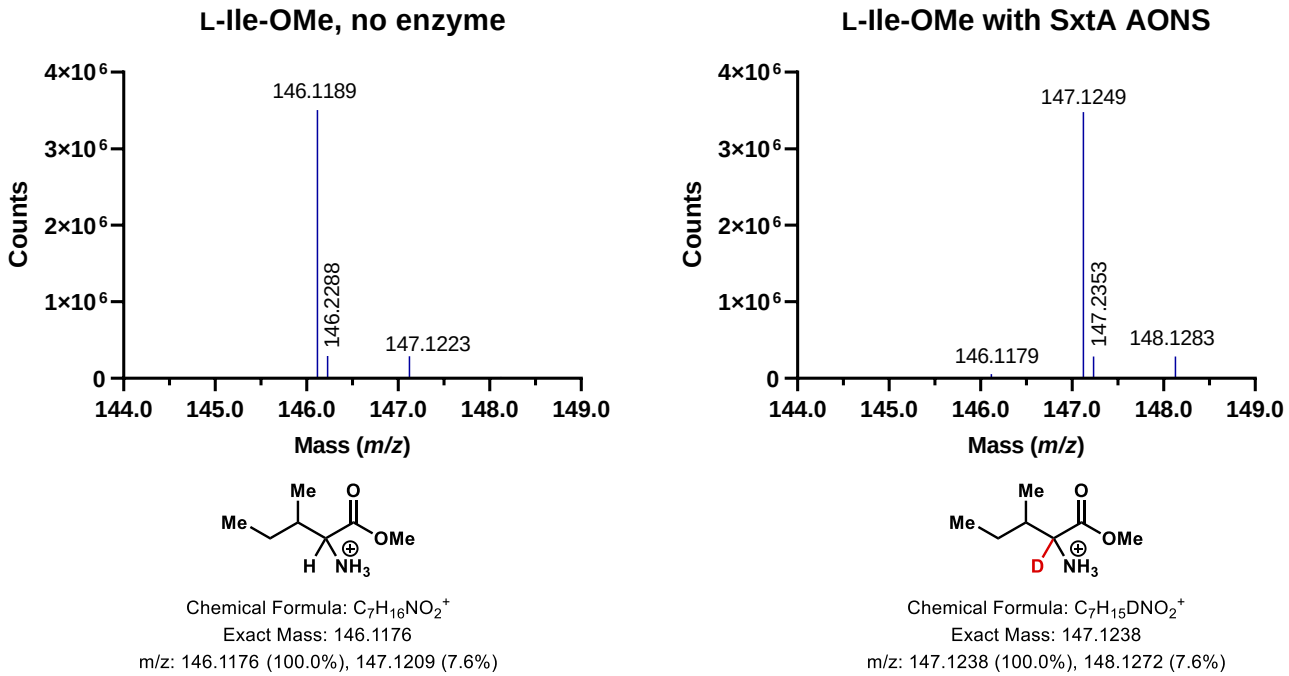
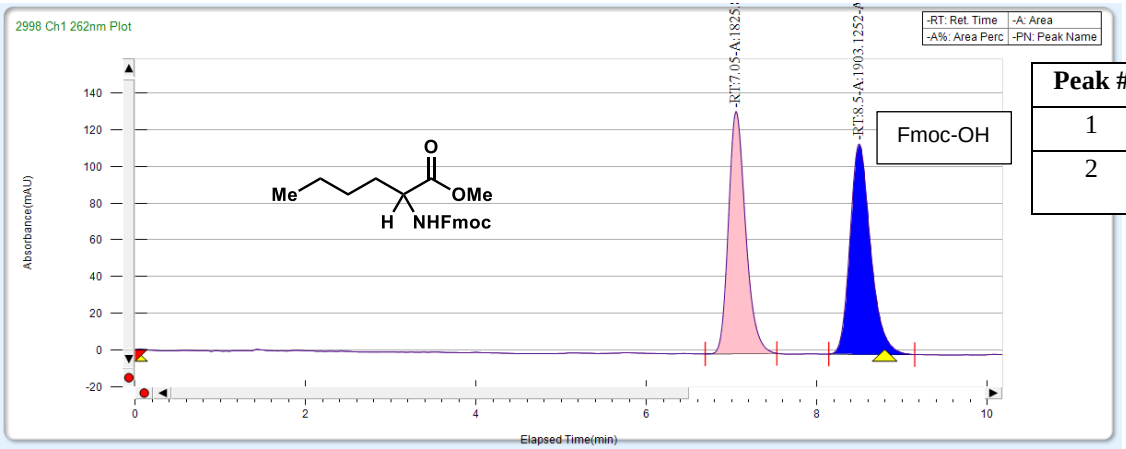


Figure 4.S10. Spectra of Ile methyl esters.
(A) PDA traces of Fmoc-protected diastereomeric isoleucine methyl esters from commercial and synthetic standards, and protected ester after incubation with SxtA AONS (CHIRALCEL OD-H, 15% MeOH with 0.5% formic acid, 3.5 mL/min); (B) representative mass spectra.

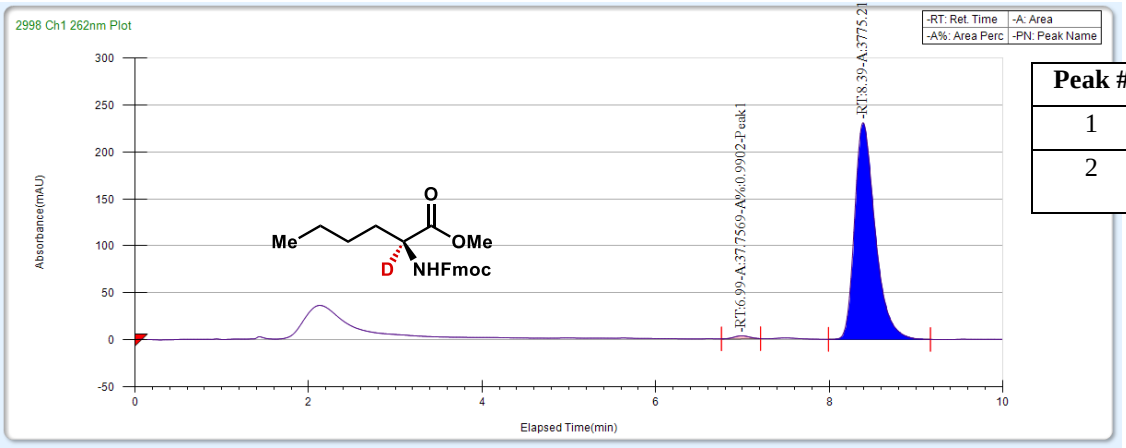
A

Racemic mixture of Nle-OMe



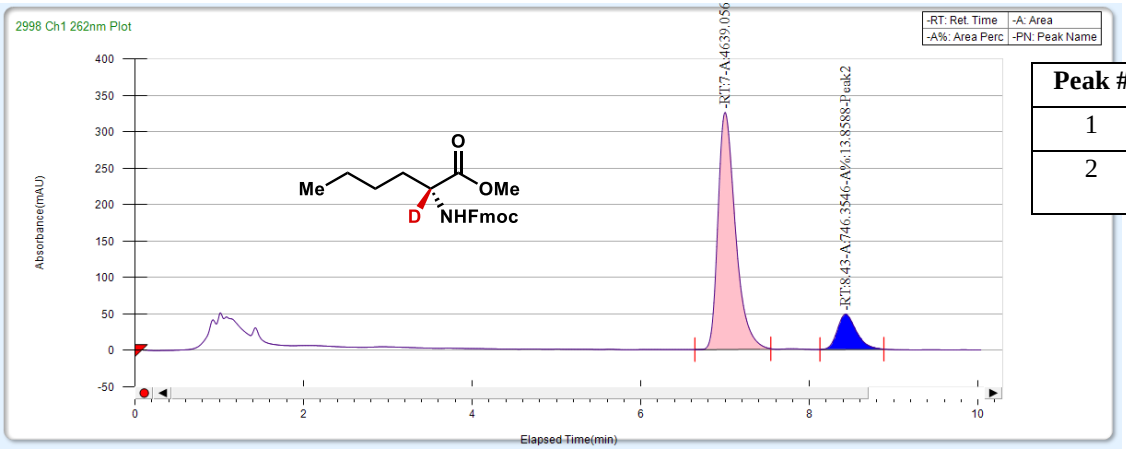
Peak #	Ret. Time	Compound	Area (%)
1	7.0	D-Nle-OMe	48.9639
2	8.5	L-Nle-OMe/ Fmoc-OH	51.0361

L-Nle-OMe incubated with SxtA AONS in D_2O



Peak #	Ret. Time	Compound	Area (%)
1	7.0	D-Nle-OMe	0.9902
2	8.4	L-Nle-OMe/ Fmoc-OH	99.0098

D-Nle-OMe incubated with SxtA AONS in D_2O



Peak #	Ret. Time	Compound	Area (%)
1	7.0	D-Nle-OMe	86.1412
2	8.4	L-Nle-OMe/ Fmoc-OH	13.8588

Due to the overlap of Fmoc-OH and L-norleucine methyl ester, these data are not included in Figure 4.10B.

B

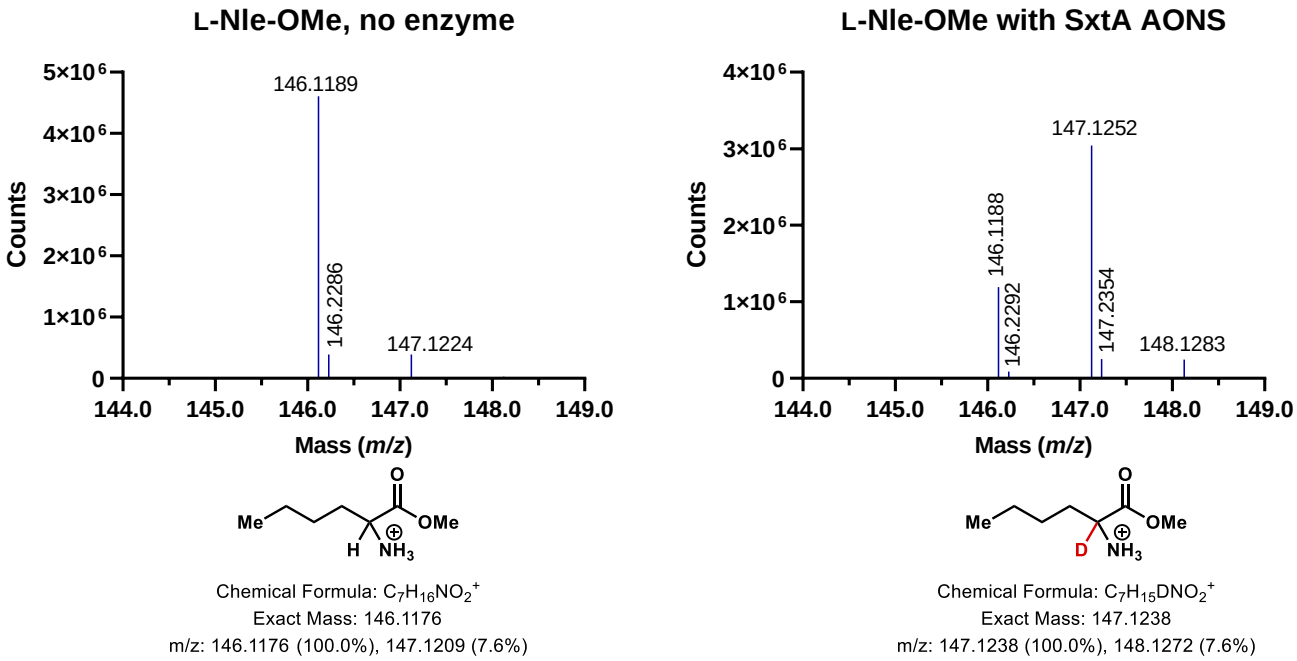
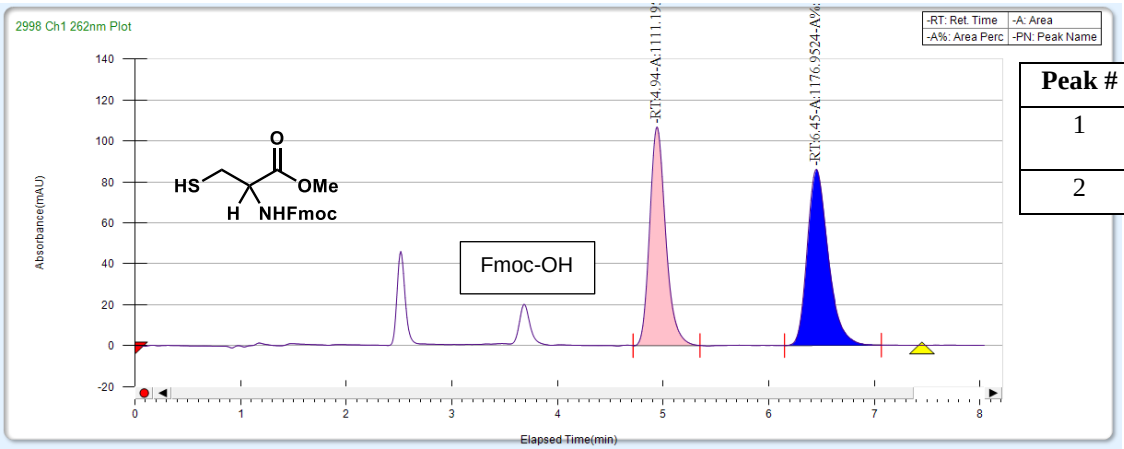


Figure 4.S11. Spectra of norleucine methyl esters.
(A) PDA traces of Fmoc-protected DL-norleucine methyl ester from commercial standards, and protected ester after incubation with SxtA AONS (CHIRALCEL OD-H, 12% MeOH with 0.5% formic acid, 3.5 mL/min); (B) representative mass spectra.

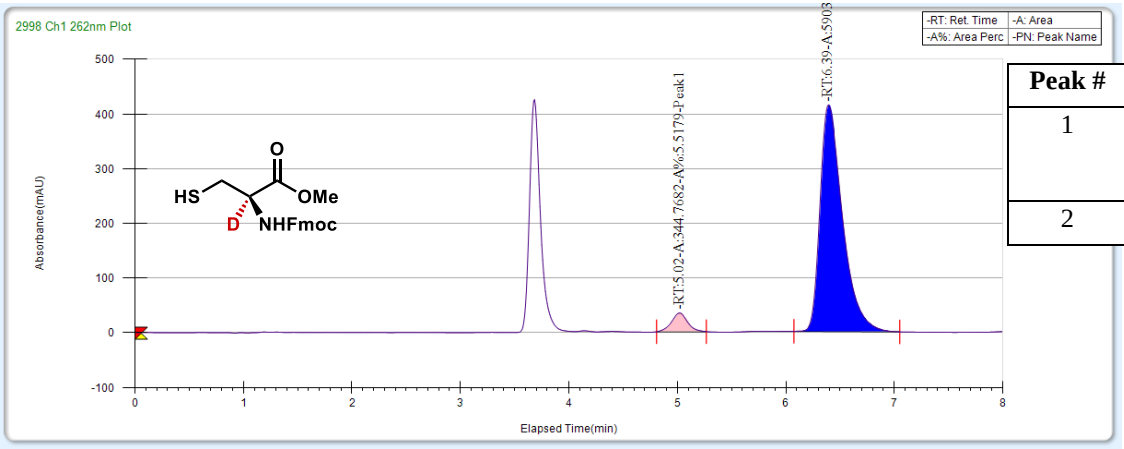
A

Racemic mixture of Cys-OMe



Peak #	Ret. Time	Compound	Area (%)
1	4.9	D-Cys-OMe w/ contaminant	48.5631
2	6.4	L-Cys-OMe	51.4369

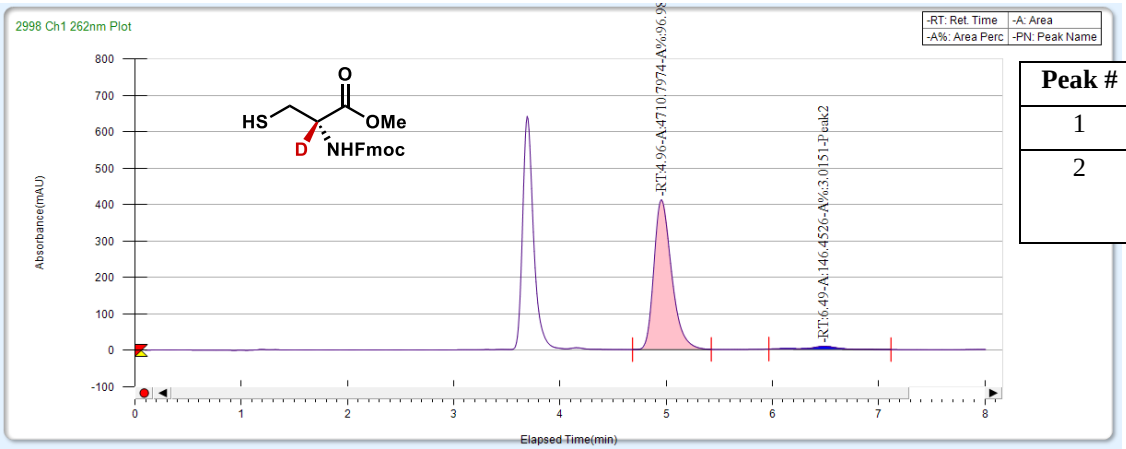
L-Cys-OMe incubated with SxtA AONS in D_2O



Peak #	Ret. Time	Compound	Area (%)
1	5.0	D-Cys-OMe w/ contaminant in negative control	5.5179
2	6.4	L-Cys-OMe	94.4821

(no Cys-OMe peaks could be seen until the addition of 100 mM TCEP)

D-Cys-OMe incubated with SxtA AONS in D_2O



Peak #	Ret. Time	Compound	Area (%)
1	4.9	D-Cys-OMe	96.8949
2	6.5	L-Cys-OMe present in standard	3.0151

(no Cys-OMe peaks could be seen until the addition of 100 mM TCEP)

B

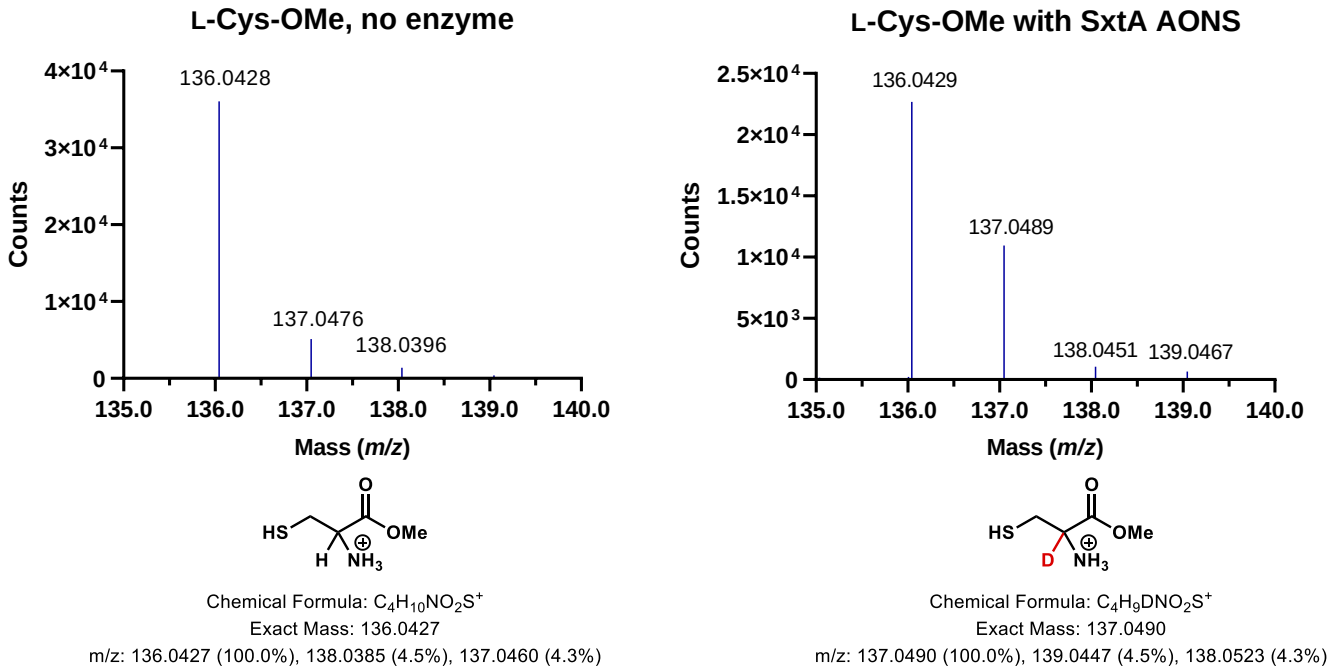
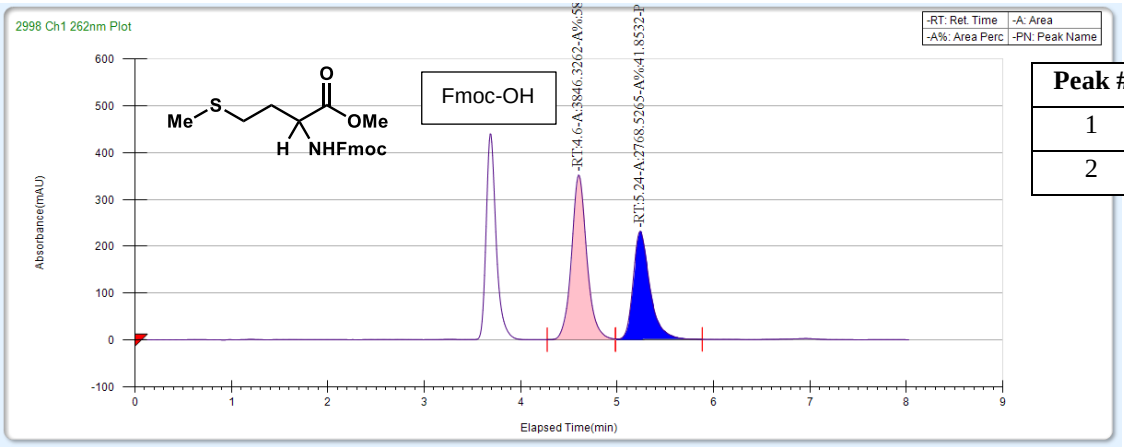


Figure 4.S12. Spectra of cysteine methyl esters.
(A) PDA traces of Fmoc-protected DL-cysteine methyl ester from commercial standards, and protected ester after incubation with SxtA AONS (CHIRALCEL OD-H, 25% MeOH with 0.5% formic acid, 3.5 mL/min); (B) representative mass spectra.

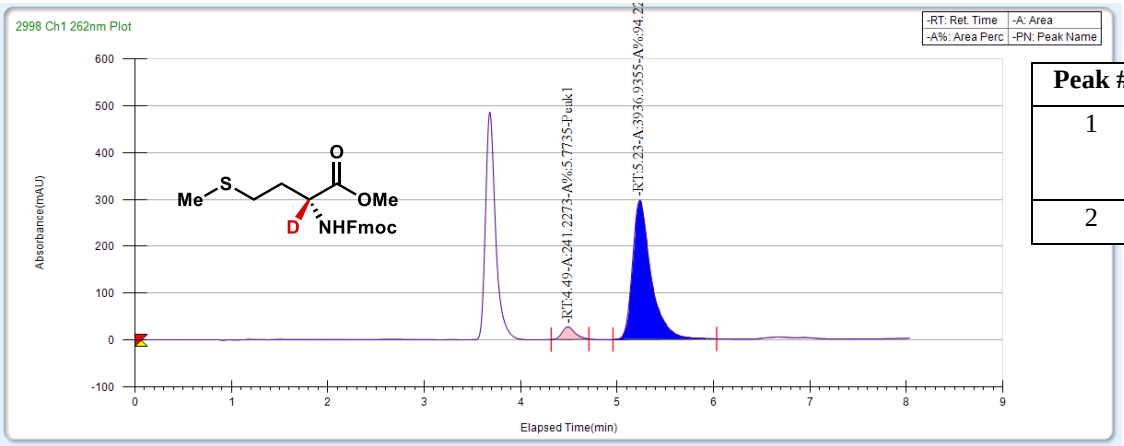
A

Racemic mixture of Met-OMe



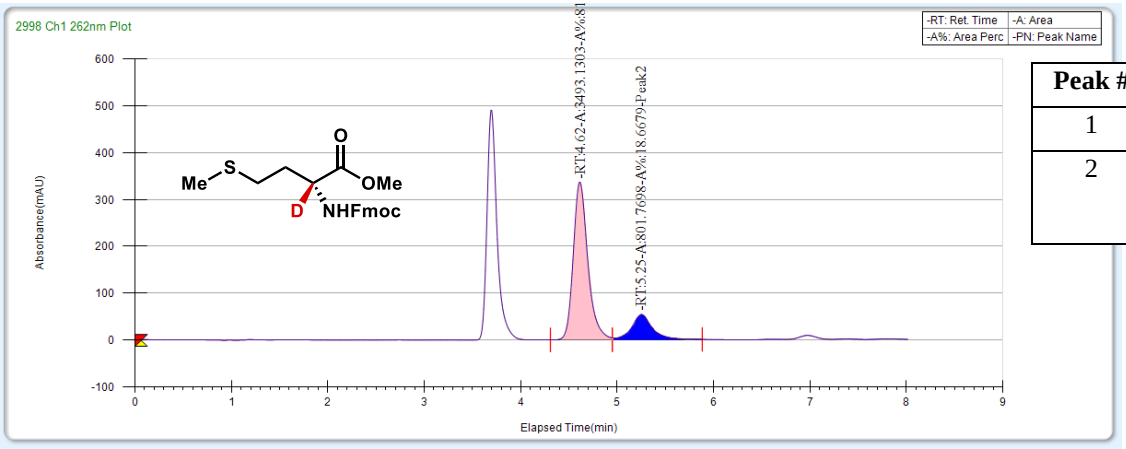
Peak #	Ret. Time	Compound	Area (%)
1	4.6	D-Met-OMe	58.1468
2	5.2	L-Met-OMe	41.8532

L-Met-OMe incubated with SxtA AONS in D_2O



Peak #	Ret. Time	Compound	Area (%)
1	4.5	D-Met-OMe present in standard	5.7735
2	5.2	L-Met-OMe	94.2265

D-Met-OMe incubated with SxtA AONS in D_2O



Peak #	Ret. Time	Compound	Area (%)
1	4.6	D-Met-OMe	81.3321
2	5.2	some L-Met-OMe present in standard	18.6679

B

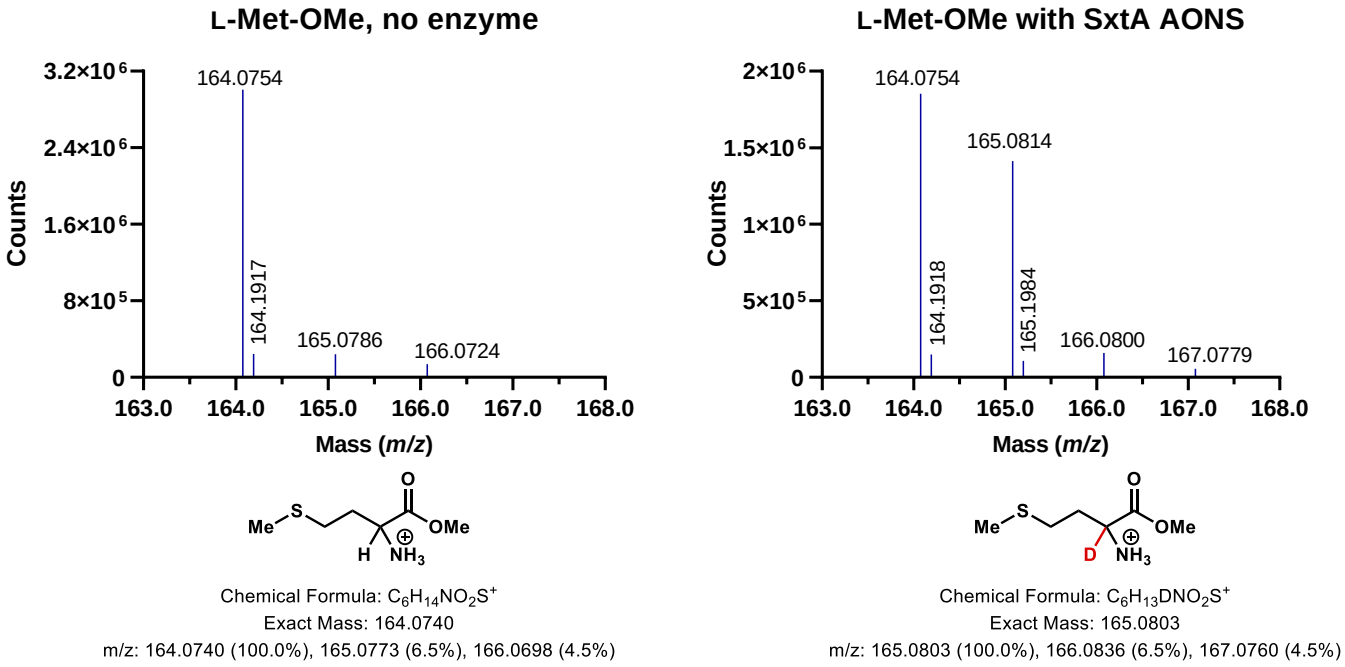
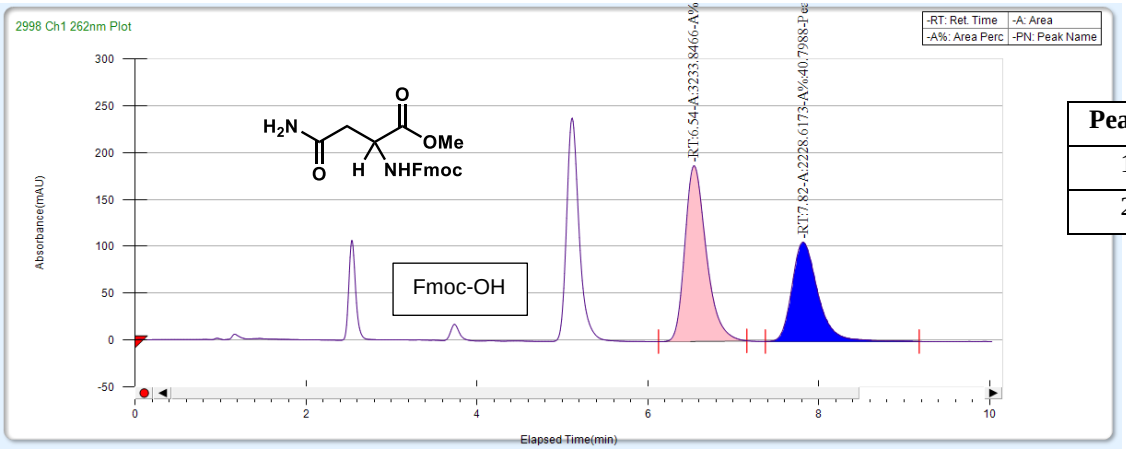


Figure 4.S13. Spectra of methionine methyl esters.
(A) PDA traces of Fmoc-protected DL-methionine methyl ester from commercial standards, and protected ester after incubation with SxtA AONS (CHIRALCEL OD-H, 25% MeOH with 0.5% formic acid, 3.5 mL/min); (B) representative mass spectra.

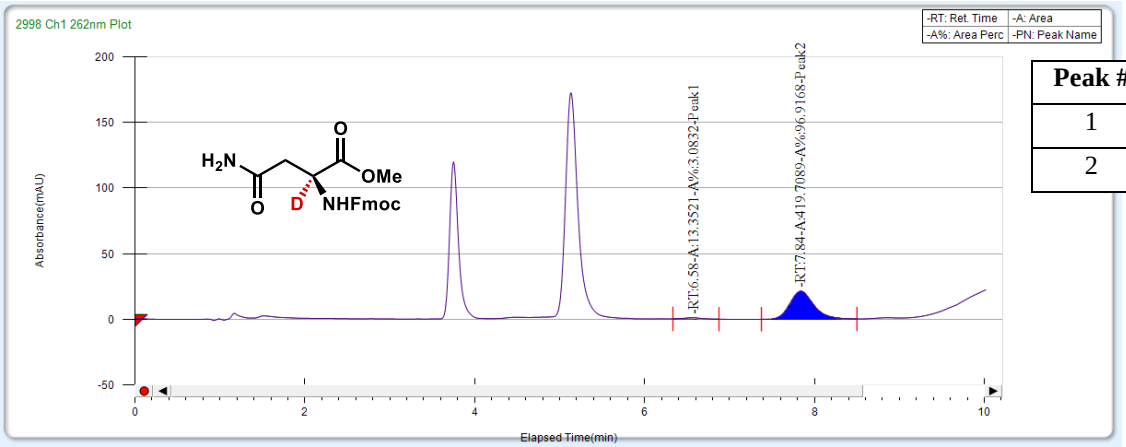
A

Racemic mixture of Asn-OMe



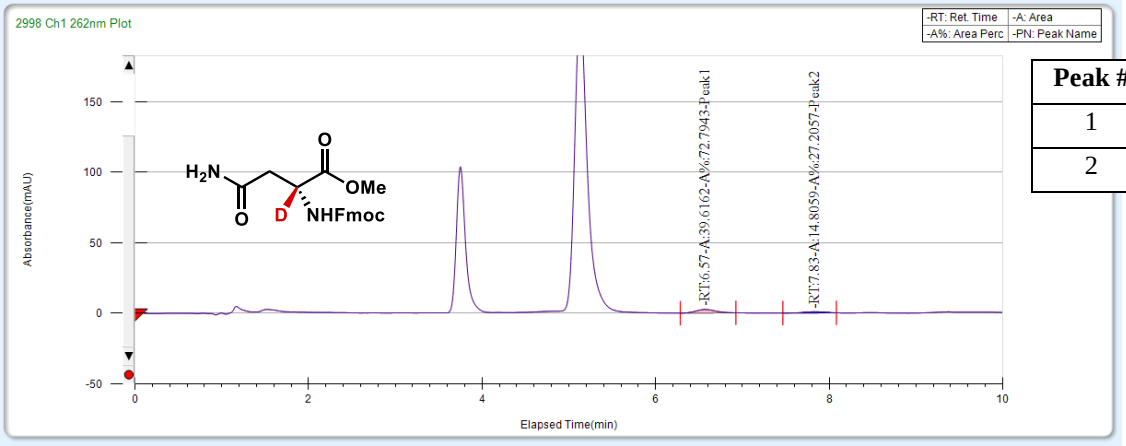
Peak #	Ret. Time	Compound	Area (%)
1	6.5	D-Asn-OMe	59.2012
2	7.8	L-Asn-OMe	40.7988

L-Asn-OMe incubated with SxtA AONS in D_2O



Peak #	Ret. Time	Compound	Area (%)
1	6.6	D-Asn-OMe	3.0832
2	7.8	L-Asn-OMe	96.9168

D-Asn-OMe incubated with SxtA AONS in D_2O



Peak #	Ret. Time	Compound	Area (%)
1	6.6	D-Asn-OMe	72.7943
2	7.8	L-Asn-OMe	27.2057

B

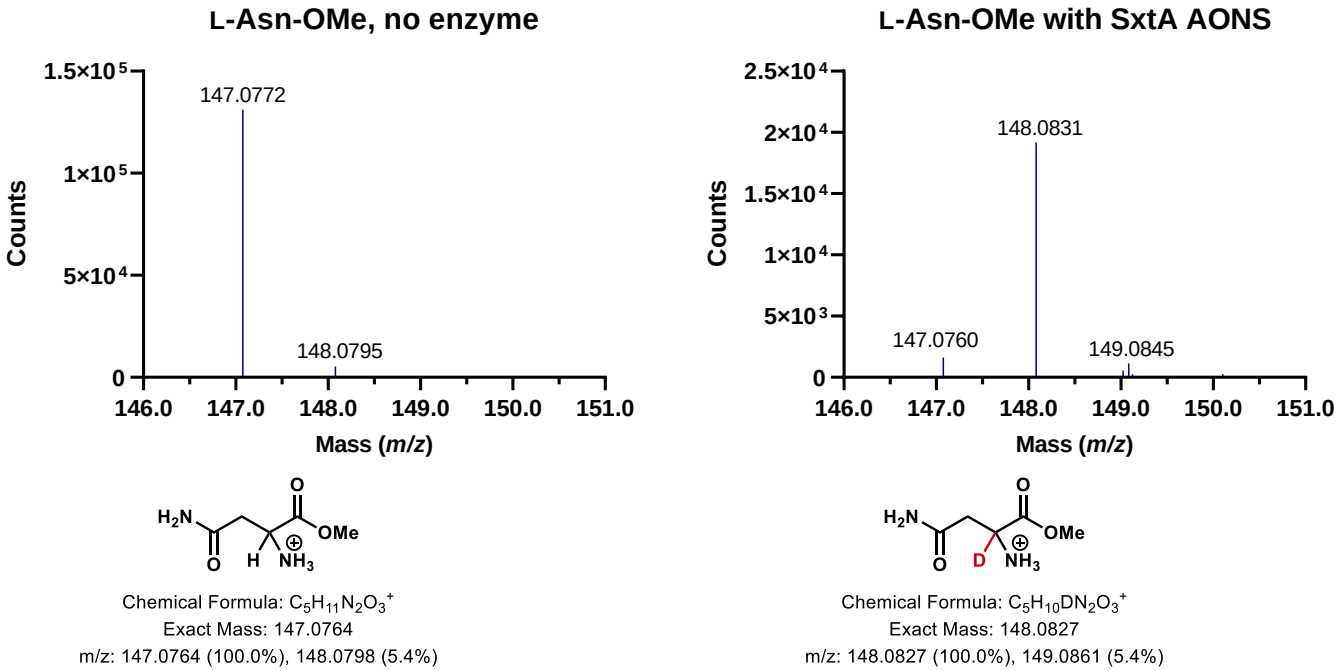
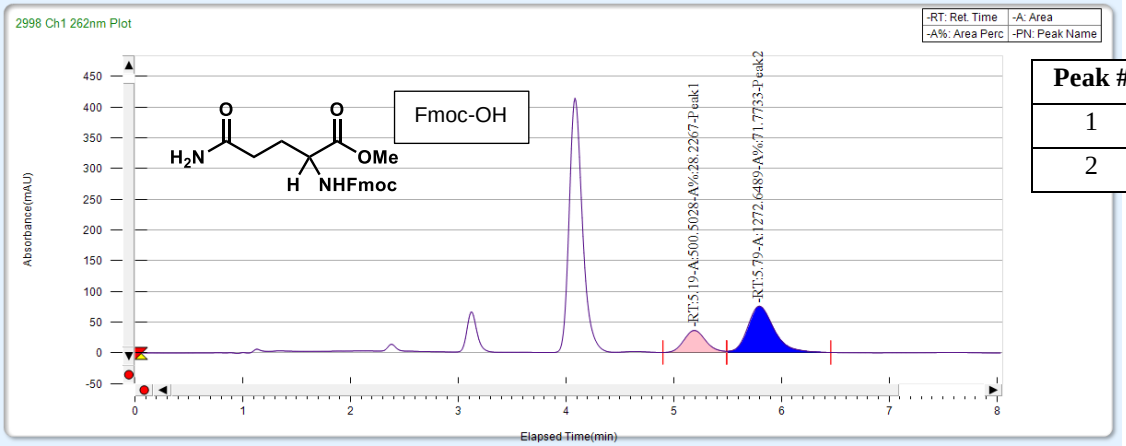


Figure 4.S14. Spectra of asparagine methyl esters.
(A) PDA traces of Fmoc-protected DL-asparagine methyl ester from synthetic standards, and protected ester after incubation with SxtA AONS (CHIRALCEL OD-H, 25% MeOH with 0.5% formic acid, 3.5 mL/min); (B) representative mass spectra.

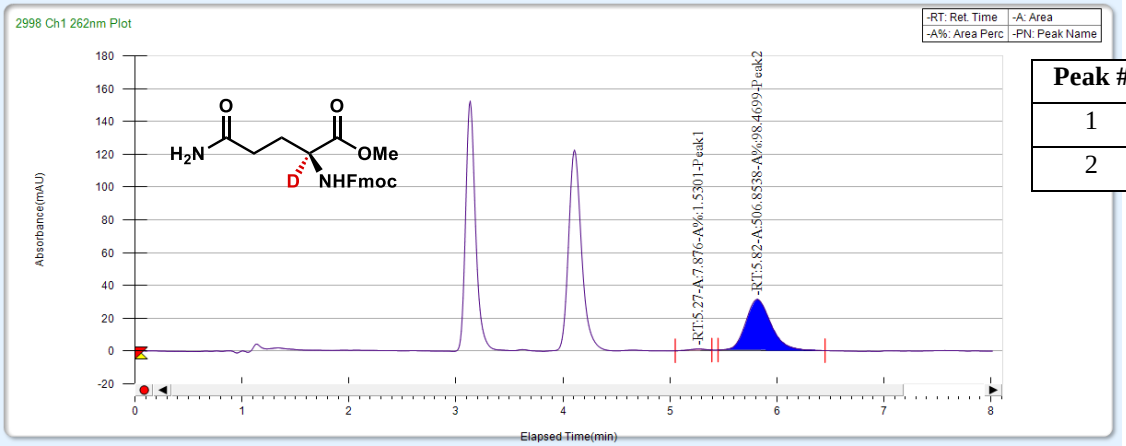
A

Racemic mixture of Gln-OMe



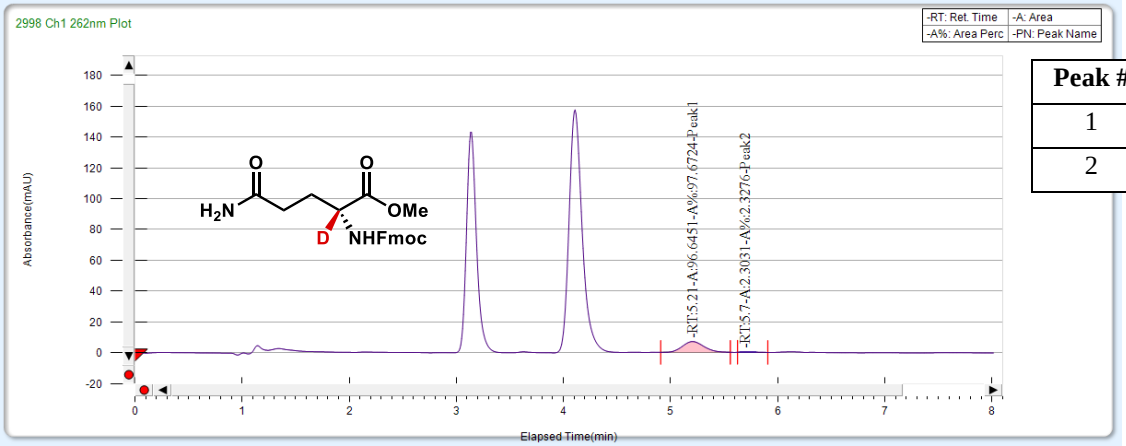
Peak #	Ret. Time	Compound	Area (%)
1	5.2	D-Gln-OMe	28.0295
2	5.8	L-Gln-OMe	71.9705

L-Gln-OMe incubated with SxtA AONS in D_2O



Peak #	Ret. Time	Compound	Area (%)
1	5.3	D-Gln-OMe	1.5301
2	5.8	L-Gln-OMe	98.4699

D-Gln-OMe incubated with SxtA AONS in D_2O



Peak #	Ret. Time	Compound	Area (%)
1	5.2	D-Gln-OMe	97.6724
2	5.8	L-Gln-OMe	2.3276

B

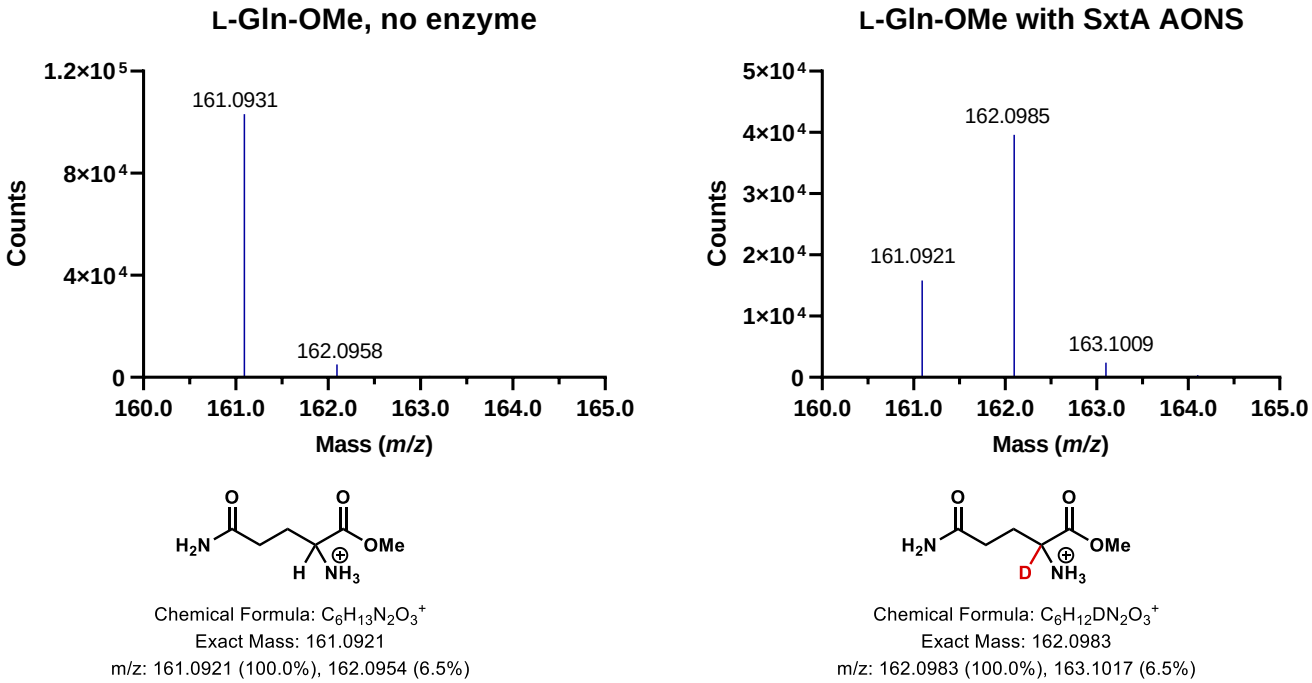
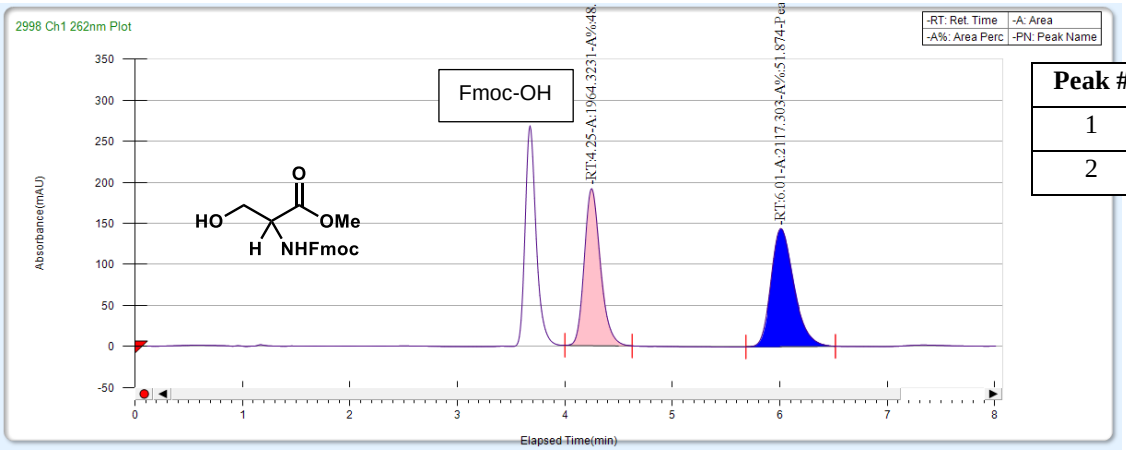


Figure 4.S15. Spectra of glutamine methyl esters.
(A) PDA traces of Fmoc-protected DL-glutamine methyl ester from synthetic standards, and protected ester after incubation with SxtA AONS (CHIRALCEL OD-H, 30% MeOH with 0.5% formic acid, 3.5 mL/min); (B) representative mass spectra.

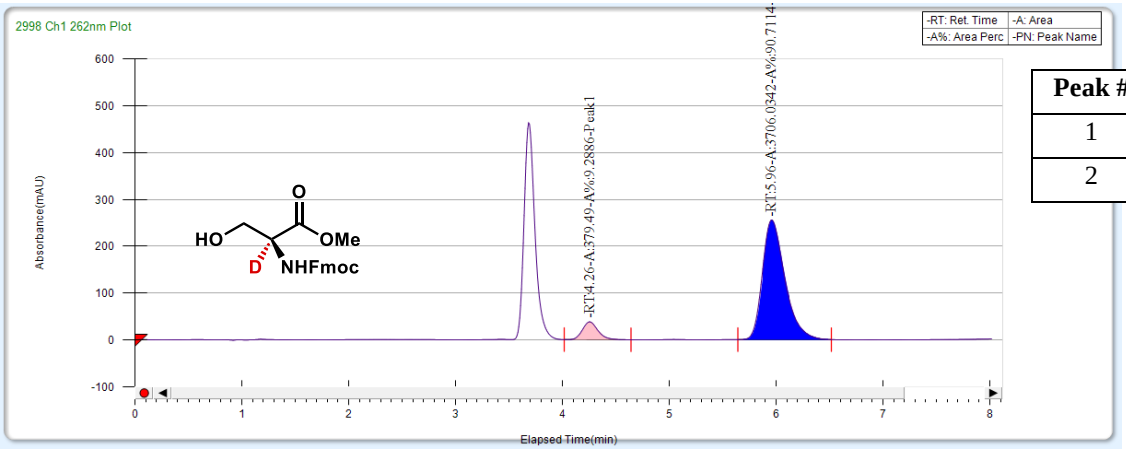
A

Racemic mixture of Ser-OMe



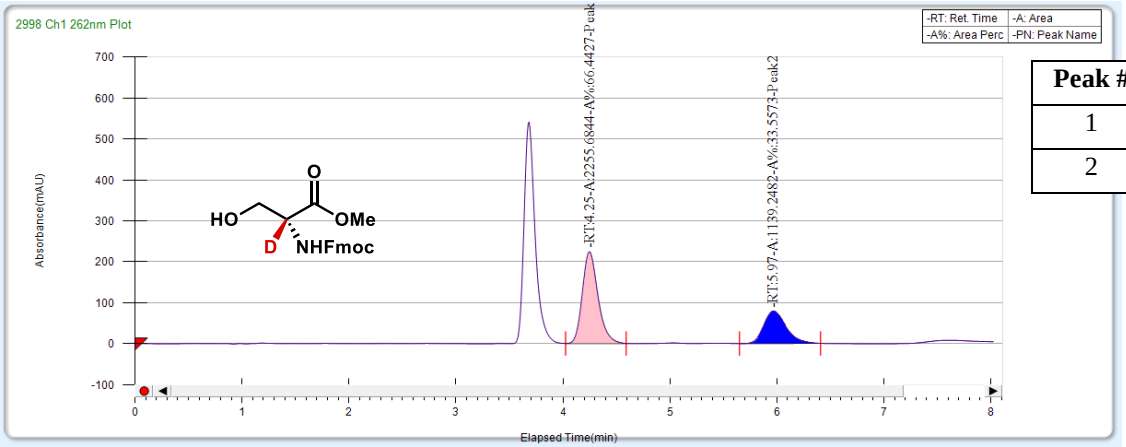
Peak #	Ret. Time	Compound	Area (%)
1	4.3	D-Ser-OMe	48.126
2	6.0	L-Ser-OMe	51.874

L-Ser-OMe incubated with SxtA AONS in D_2O



Peak #	Ret. Time	Compound	Area (%)
1	4.3	D-Ser-OMe	9.2886
2	6.0	L-Ser-OMe	90.7114

D-Ser-OMe incubated with SxtA AONS in D_2O



Peak #	Ret. Time	Compound	Area (%)
1	4.3	D-Ser-OMe	66.4427
2	6.0	L-Ser-OMe	33.5573

B

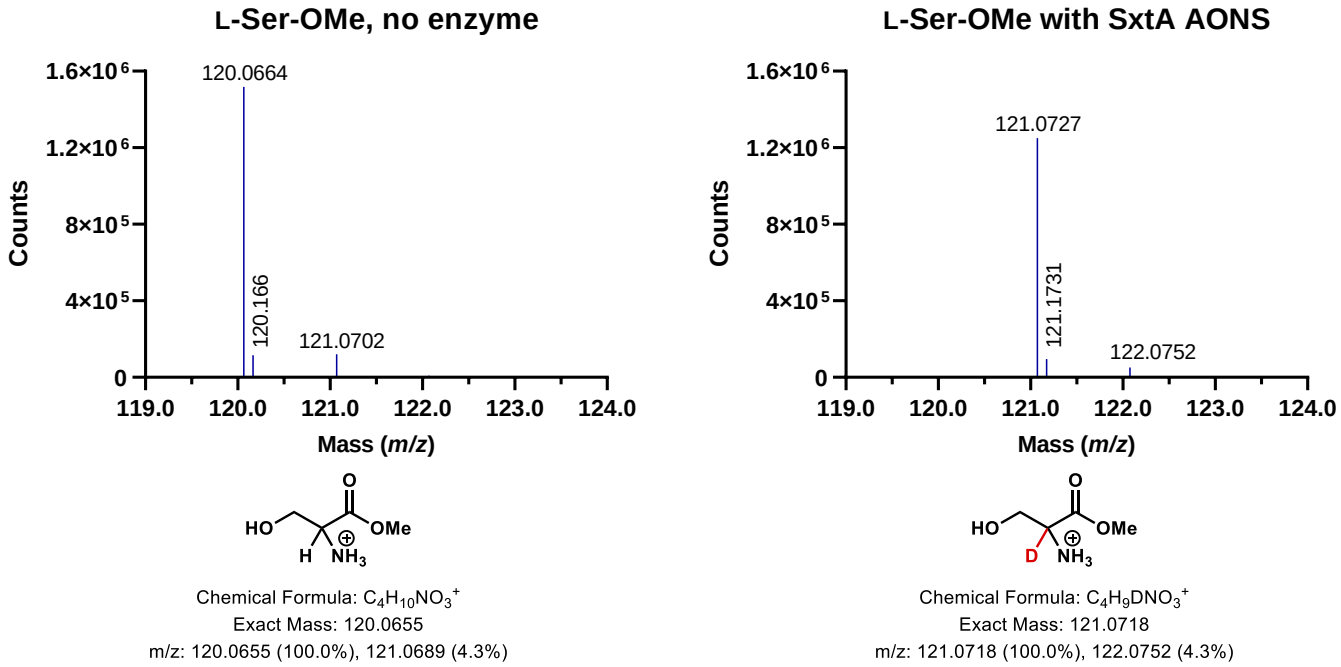
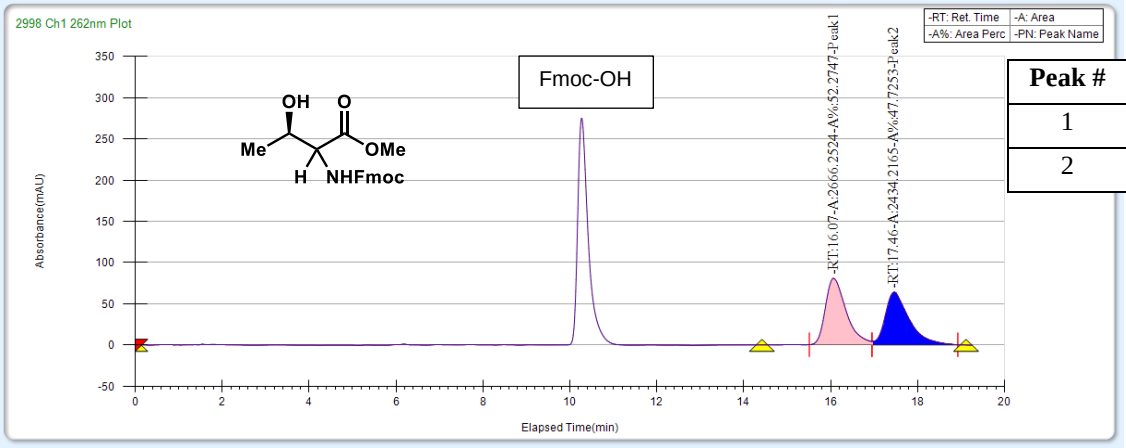


Figure 4.S16. Spectra of serine methyl esters.
(A) PDA traces of Fmoc-protected DL-serine methyl ester from commercial standards, and protected ester after incubation with SxtA AONS (CHIRALCEL OD-H, 30% MeOH with 0.5% formic acid, 3.5 mL/min); (B) representative mass spectra.

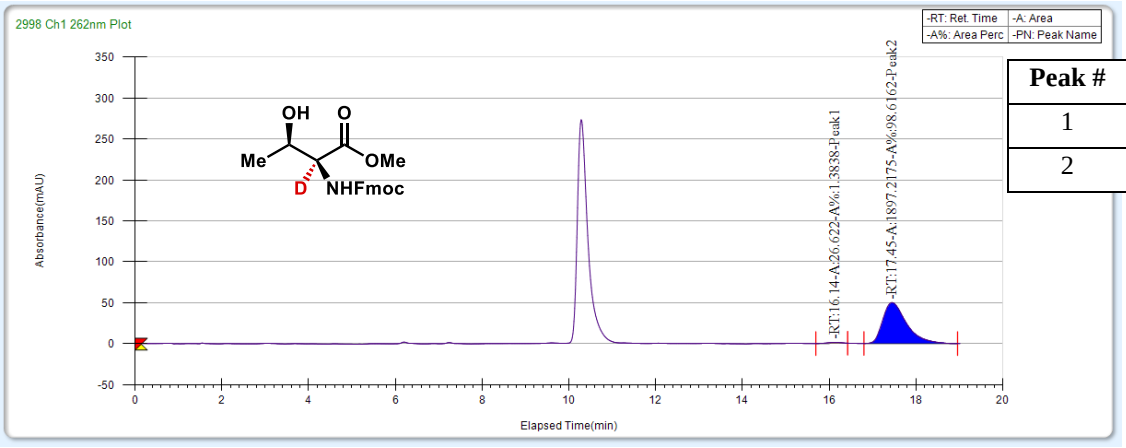
A

Diastereomeric mixture of L-Thr-OMe and D-*allo*-Thr-OMe



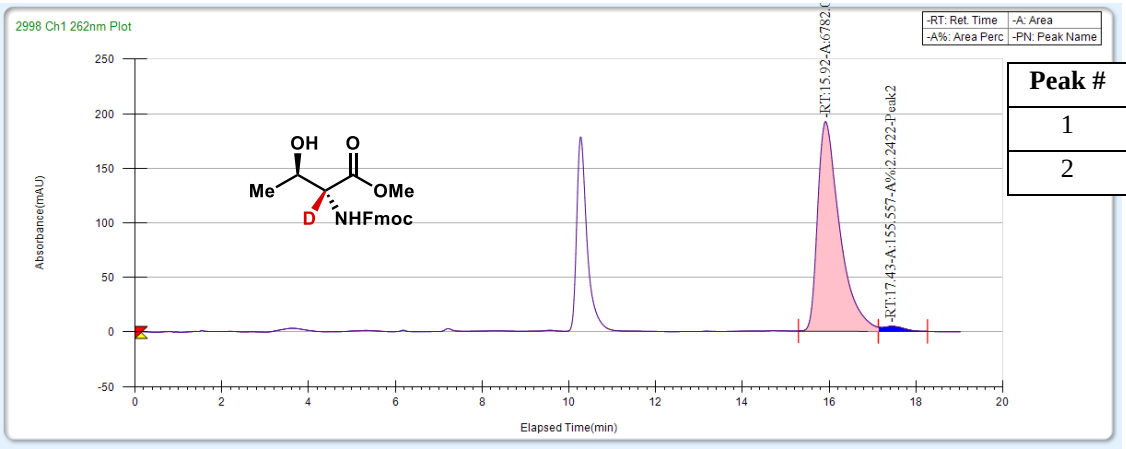
Peak #	Ret. Time	Compound	Area (%)
1	16.1	D-*allo*-Thr-OMe	52.2747
2	17.5	L-Thr-OMe	47.7253

L-Thr-OMe incubated with SxtA AONS in D_2O



Peak #	Ret. Time	Compound	Area (%)
1	16.1	D-*allo*-Thr-OMe	1.3838
2	17.5	L-Thr-OMe	98.6162

D-*allo*-Thr-OMe incubated with SxtA AONS in D_2O



Peak #	Ret. Time	Compound	Area (%)
1	16.1	D-*allo*-Thr-OMe	97.7578
2	17.4	L-Thr-OMe	2.2422

B

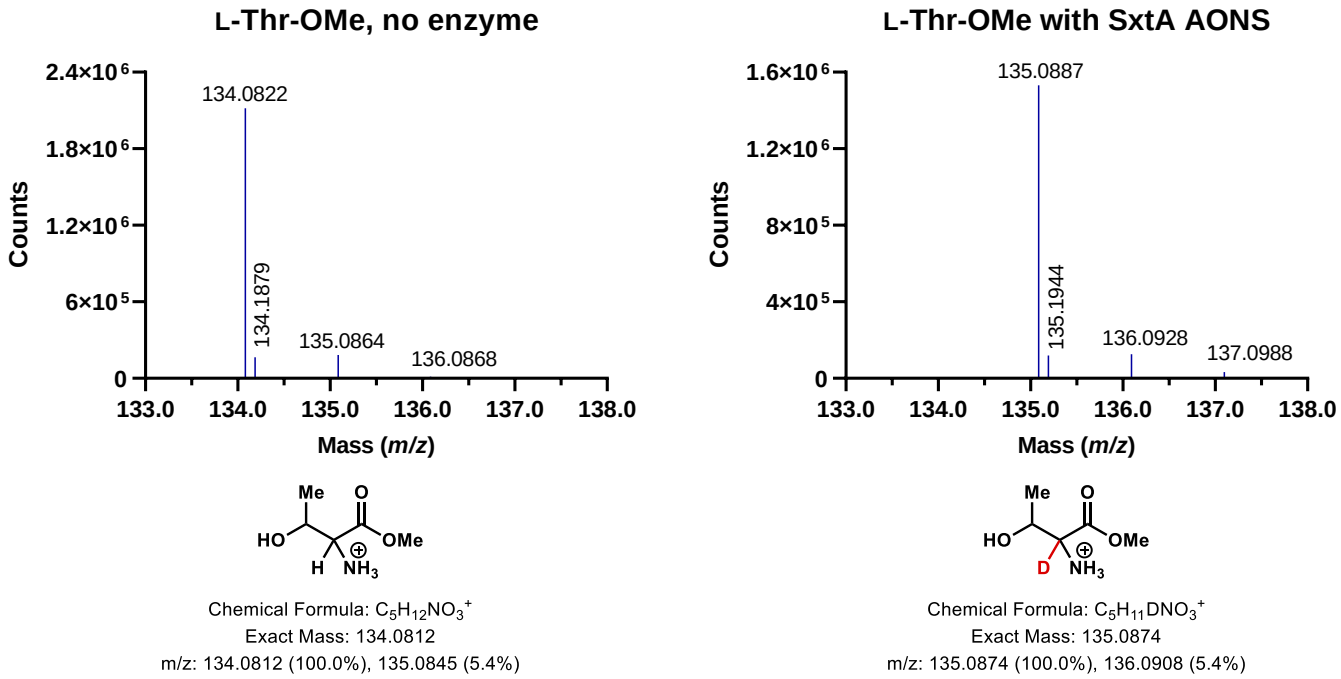
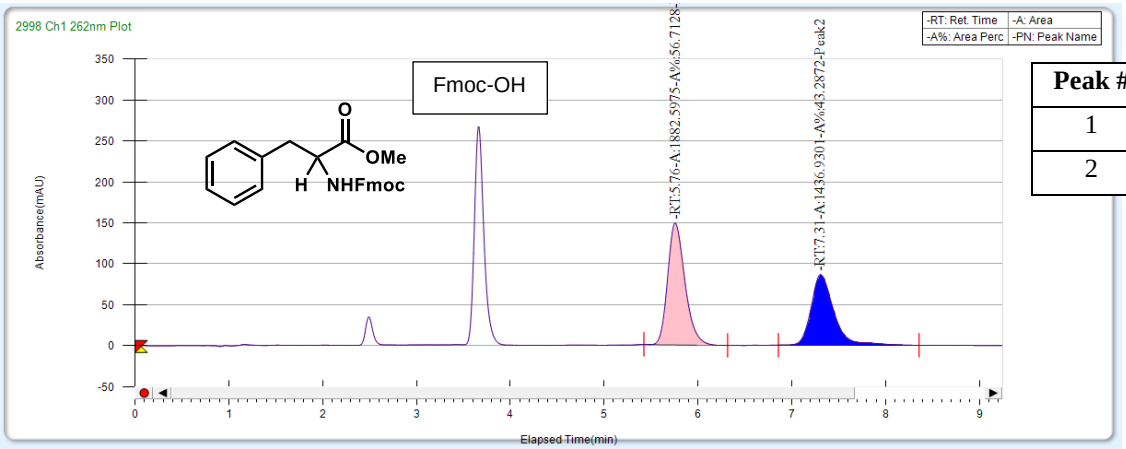


Figure 4.S17. Spectra of threonine methyl esters.
(A) PDA traces of Fmoc-protected diastereomeric mixture of L-threonine methyl ester and D-*allo*-Thr-OMe derivatized from commercial and synthetic standards, and protected ester after incubation with SxtA AONS (CHIRALCEL OD-H, 10% MeOH with 0.5% formic acid, 3.5 mL/min); (B) representative mass spectra.

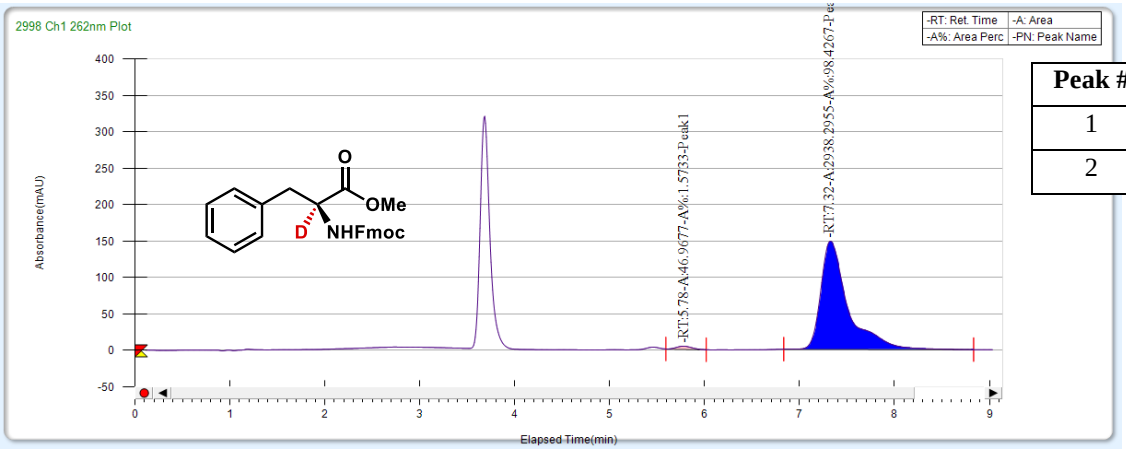
A

Racemic mixture of Phe-OMe



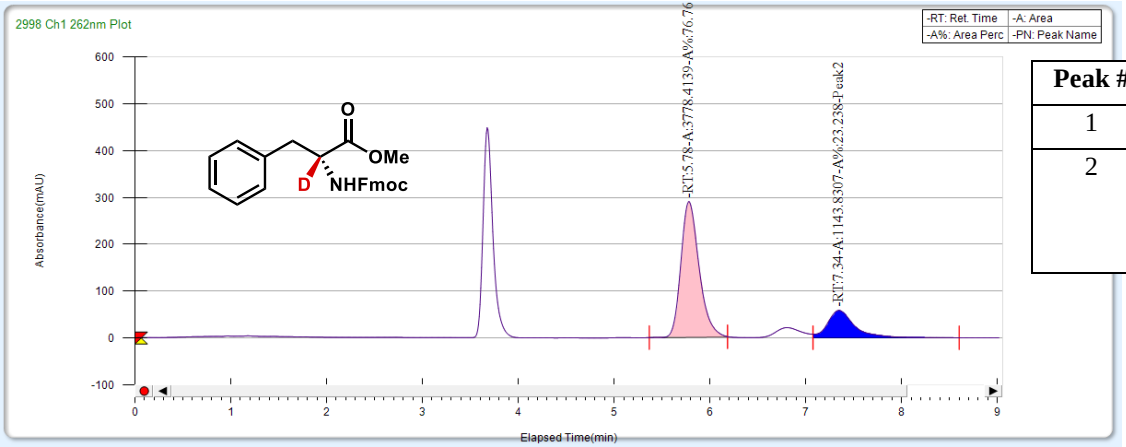
Peak #	Ret. Time	Compound	Area (%)
1	5.8	D-Phe-OMe	56.7128
2	7.3	L-Phe-OMe	43.2872

L-Phe-OMe incubated with SxtA AONS in D_2O



Peak #	Ret. Time	Compound	Area (%)
1	5.8	D-Phe-OMe	1.5733
2	7.3	L-Phe-OMe	98.4267

D-Phe-OMe incubated with SxtA AONS in D_2O



Peak #	Ret. Time	Compound	Area (%)
1	5.8	D-Phe-OMe	76.762
2	7.3	L-Phe-OMe w/ some observed in negative control	23.238

B

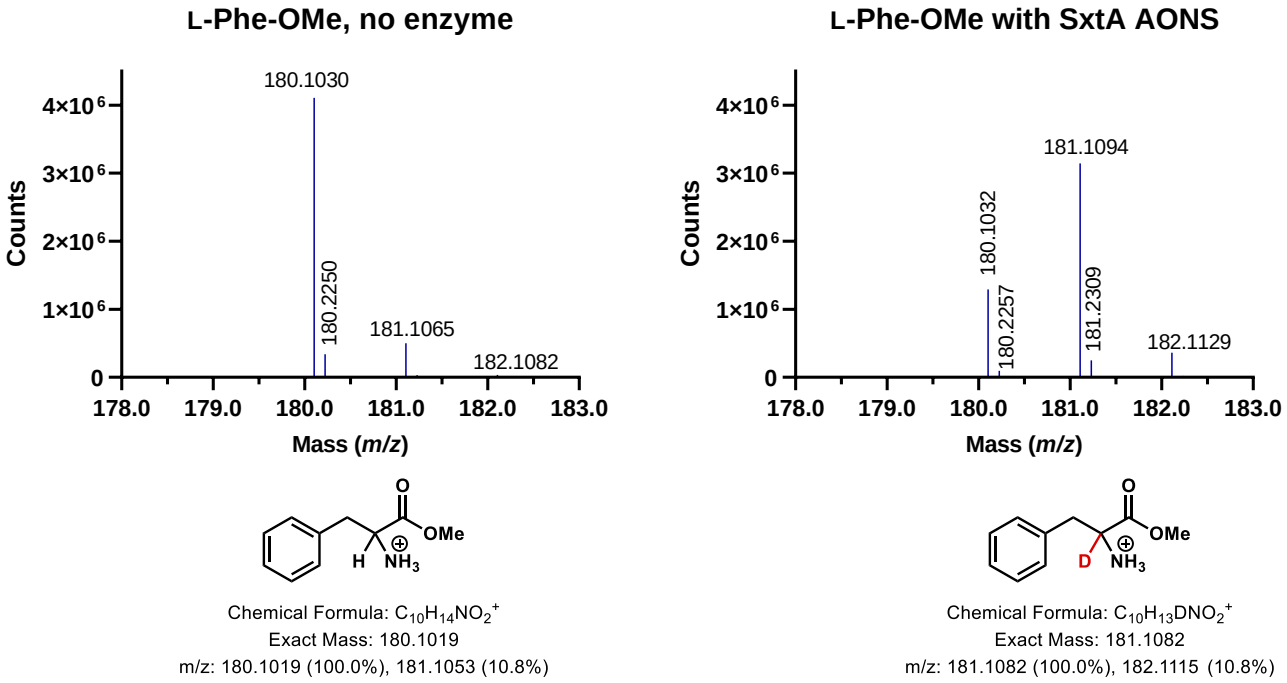
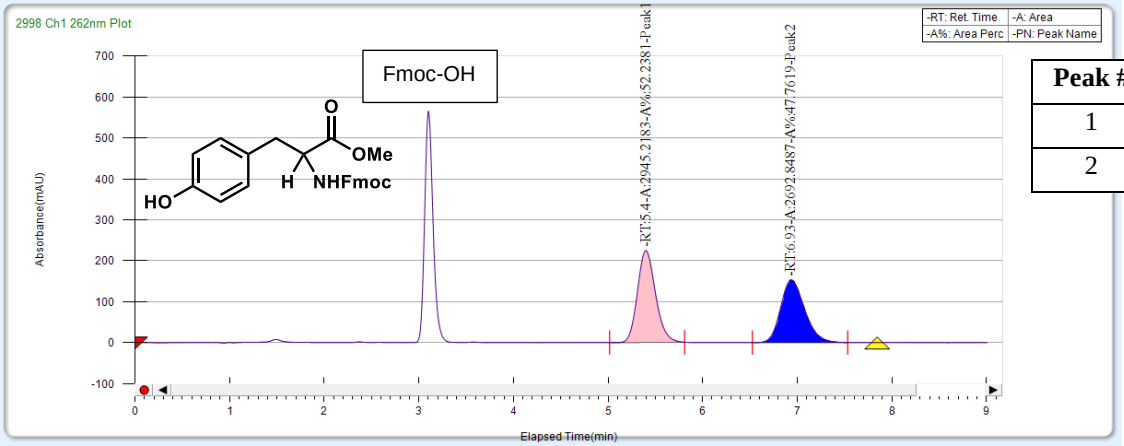


Figure 4.S18. Spectra of phenylalanine methyl esters.
(A) PDA traces of Fmoc-protected DL-phenylalanine methyl ester from commercial or synthetic standards, and protected ester after incubation with SxtA AONS (CHIRALCEL OD-H, 25% MeOH with 0.5% formic acid, 3.5 mL/min); (B) representative mass spectra.

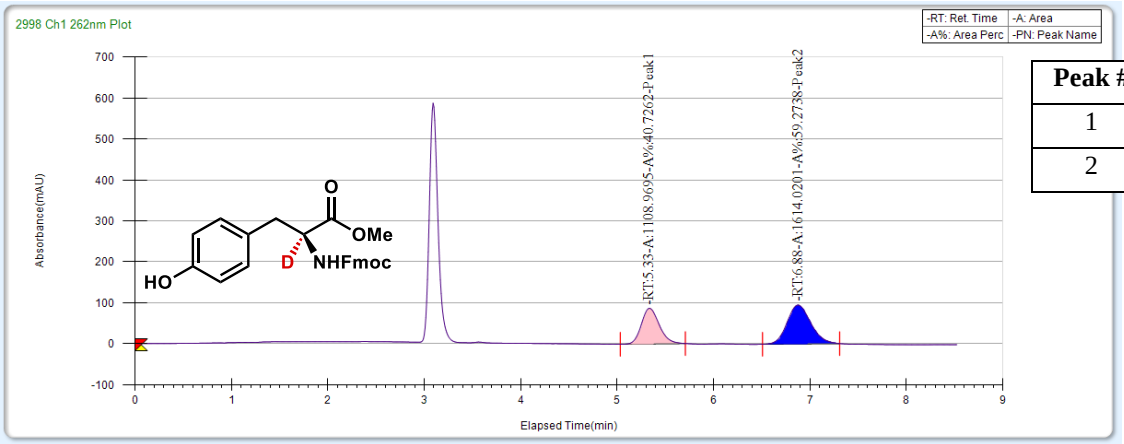
A

Racemic mixture of Tyr-OMe



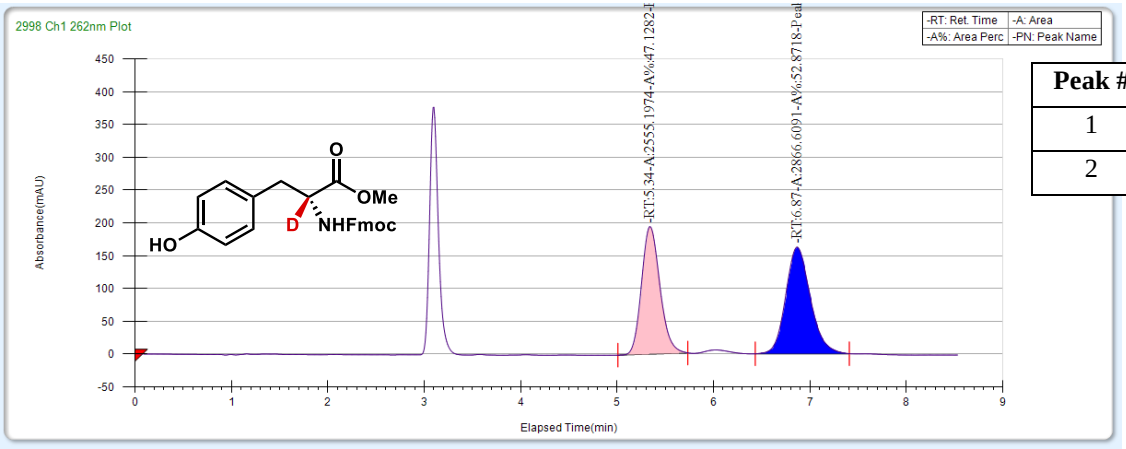
Peak #	Ret. Time	Compound	Area (%)
1	5.4	D-Tyr-OMe	52.2381
2	6.9	L-Tyr-OMe	47.7619

L-Tyr-OMe incubated with SxtA AONS in D_2O



Peak #	Ret. Time	Compound	Area (%)
1	5.3	D-Tyr-OMe	40.7262
2	6.9	L-Tyr-OMe	59.2738

D-Tyr-OMe incubated with SxtA AONS in D_2O



Peak #	Ret. Time	Compound	Area (%)
1	5.3	D-Tyr-OMe	47.1282
2	6.9	L-Tyr-OMe	52.8718

B

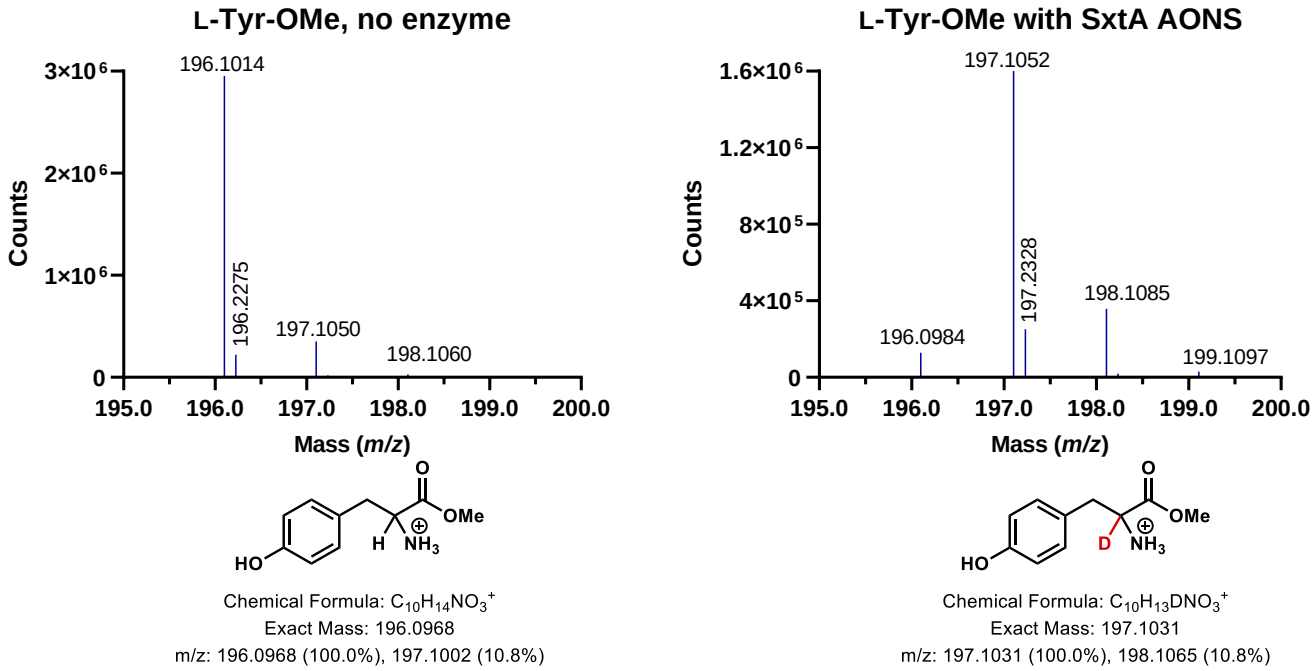
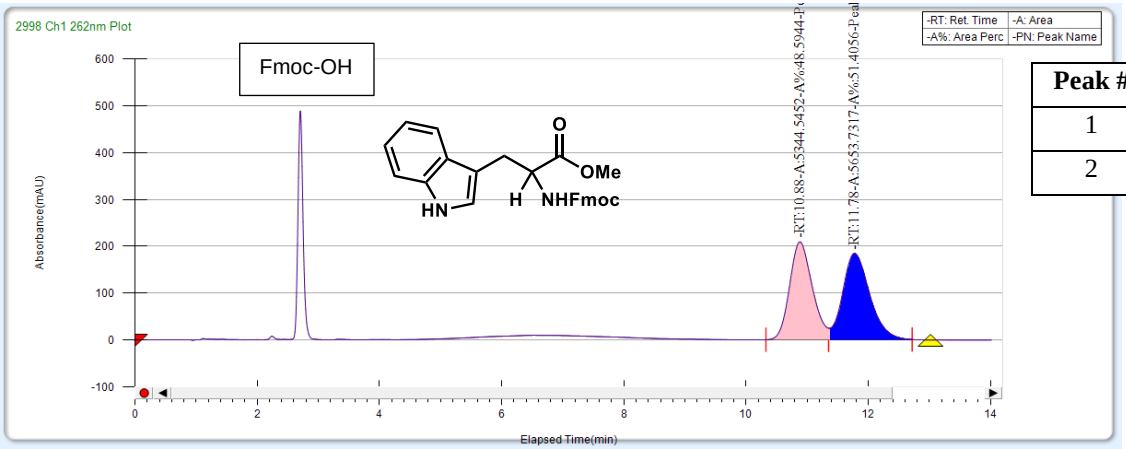


Figure 4.S19. Spectra of tyrosine methyl esters.
(A) PDA traces of Fmoc-protected r DL-tyrosine methyl ester from commercial standards, and protected ester after incubation with SxtA AONS (CHIRALCEL OD-H, 30% MeOH with 0.5% formic acid, 3.5 mL/min); (B) representative mass spectra.

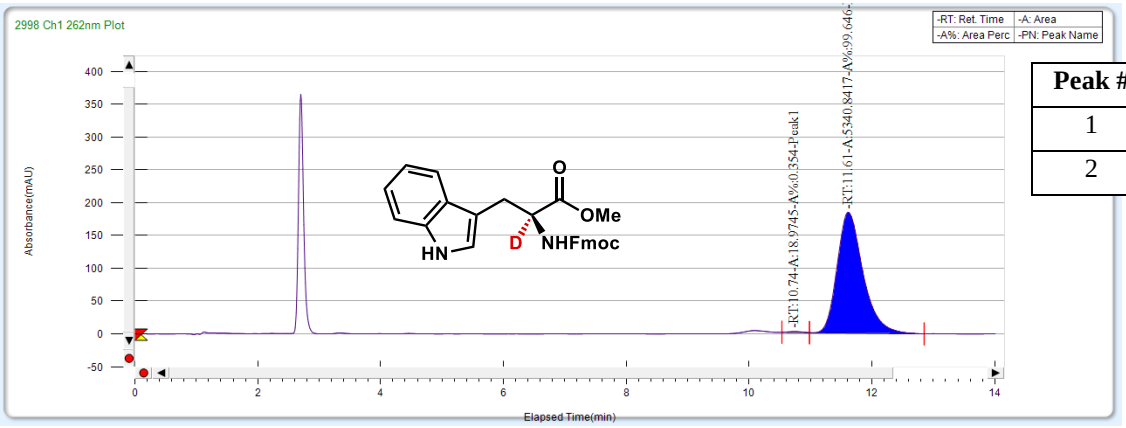
A

Racemic mixture of Trp-OMe



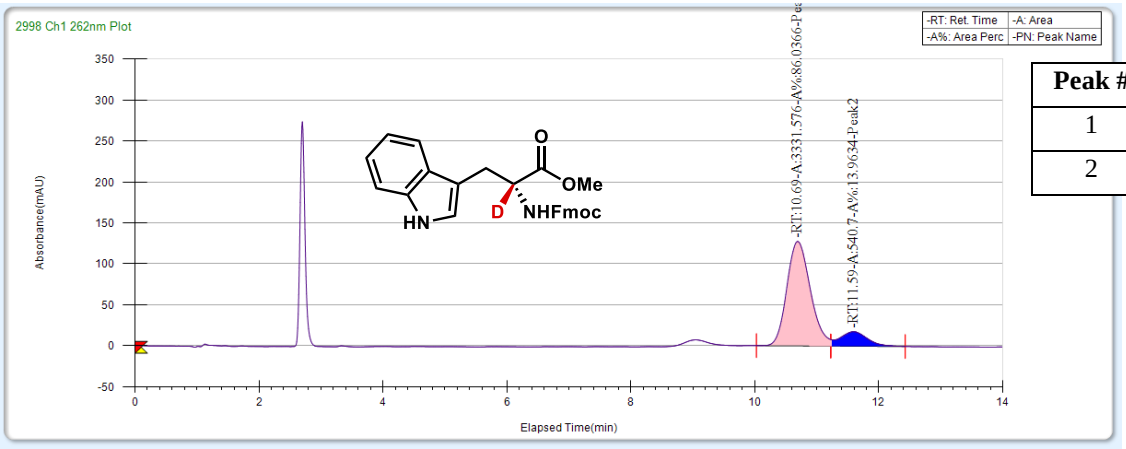
Peak #	Ret. Time	Compound	Area (%)
1	10.9	D-Trp-OMe	48.5944
2	11.8	L-Trp-OMe	51.4057

L-Trp-OMe incubated with SxtA AONS in D_2O



Peak #	Ret. Time	Compound	Area (%)
1	10.9	D-Trp-OMe	0.354
2	11.8	L-Trp-OMe	99.646

D-Trp-OMe incubated with SxtA AONS in D_2O



Peak #	Ret. Time	Compound	Area (%)
1	10.9	D-Trp-OMe	86.0366
2	11.8	L-Trp-OMe	13.9634

B

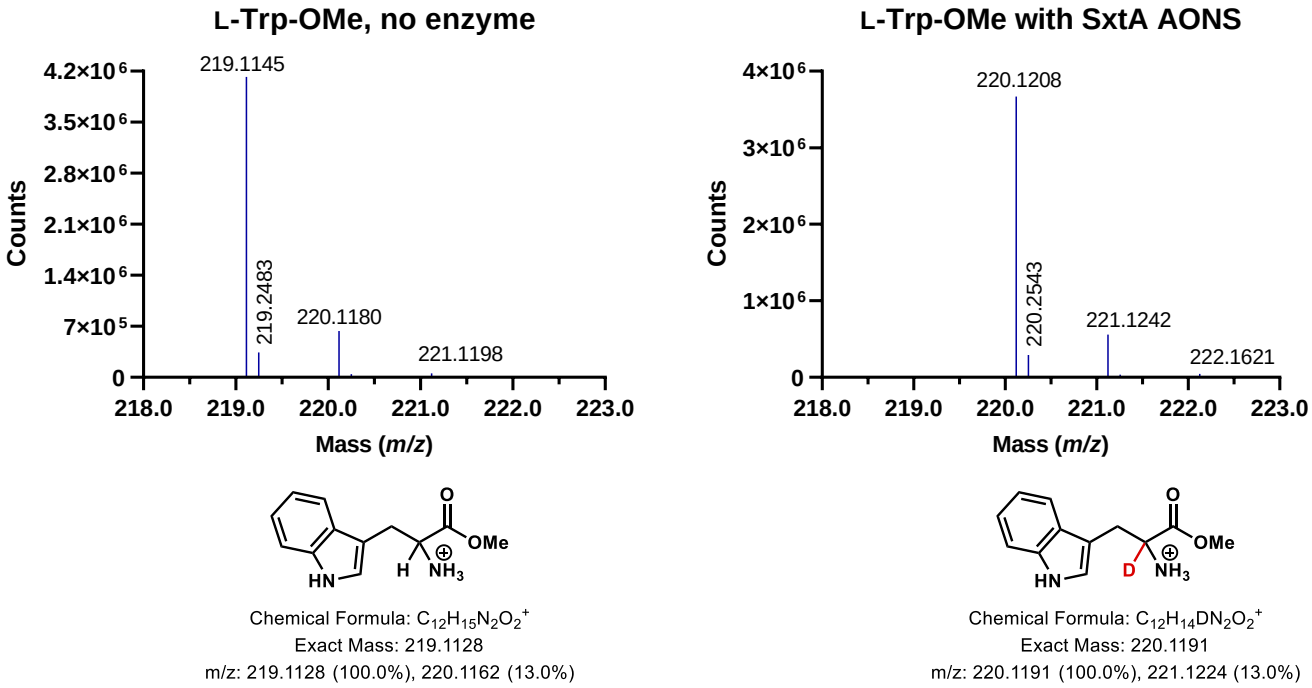
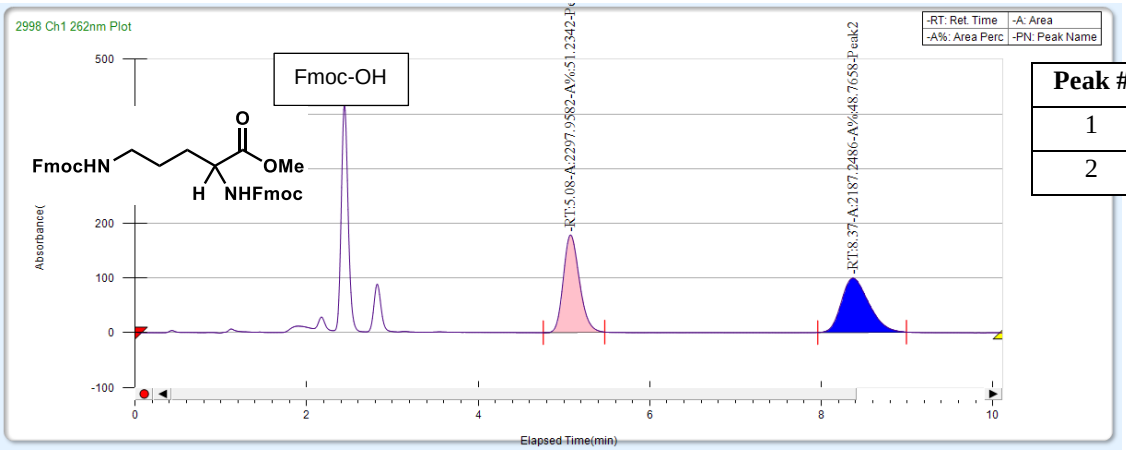


Figure 4.S20. Spectra of tryptophan methyl esters.
(A) PDA traces of Fmoc-protected DL-tryptophan methyl ester from commercial standards, and protected ester after incubation with SxtA AONS (CHIRALCEL OD-H, 35% MeOH with 0.5% formic acid, 3.5 mL/min); (B) representative mass spectra.

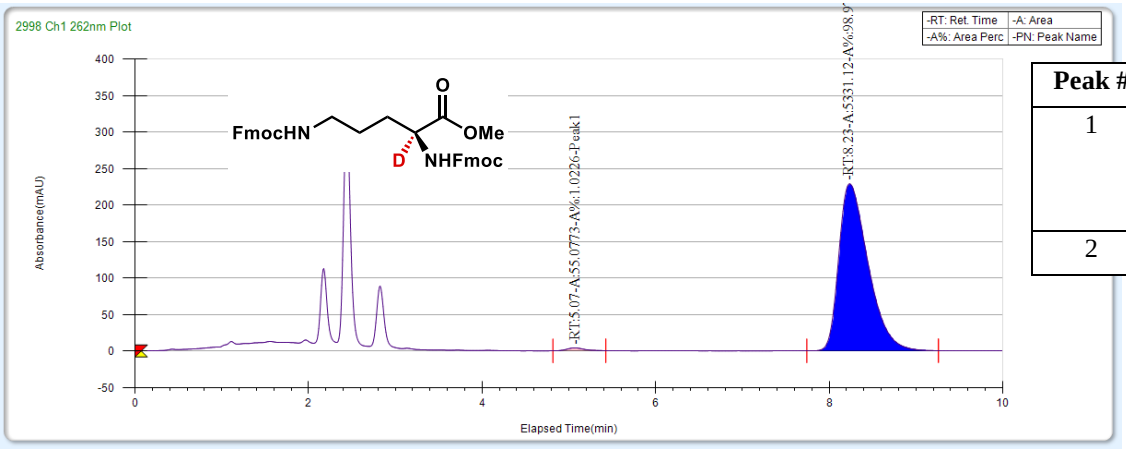
A

Racemic mixture of Orn-OMe



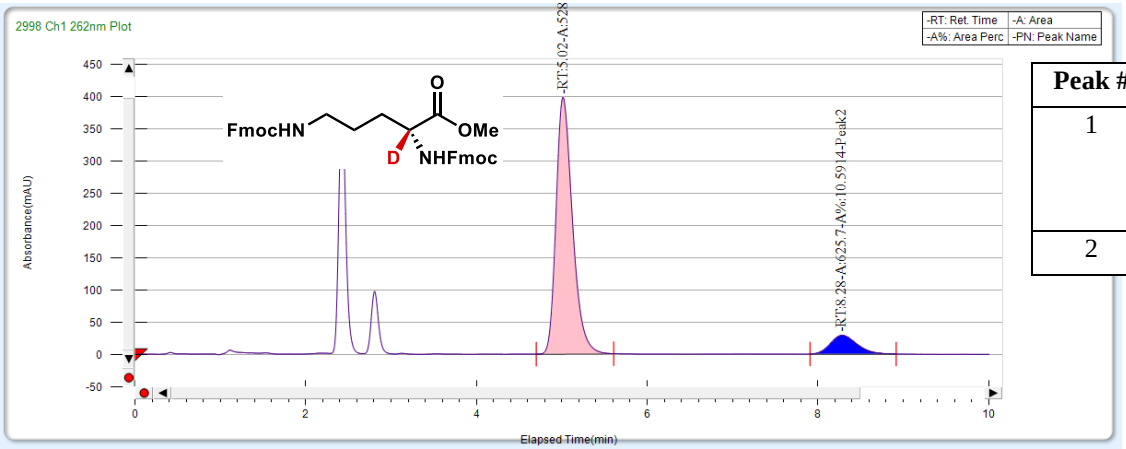
Peak #	Ret. Time	Compound	Area (%)
1	5.1	D-Orn-OMe	51.2342
2	8.4	L-Orn-OMe	48.7658

L-Orn-OMe incubated with SxtA AONS in D_2O



Peak #	Ret. Time	Compound	Area (%)
1	5.1	D-Orn-OMe w/ some observed in negative control	1.0266
2	8.2	L-Orn-OMe	98.9774

D-Orn-OMe incubated with SxtA AONS in D_2O



Peak #	Ret. Time	Compound	Area (%)
1	5.1	D-Orn-OMe w/ some observed in negative control	89.4086
2	8.2	L-Orn-OMe	10.5914

B

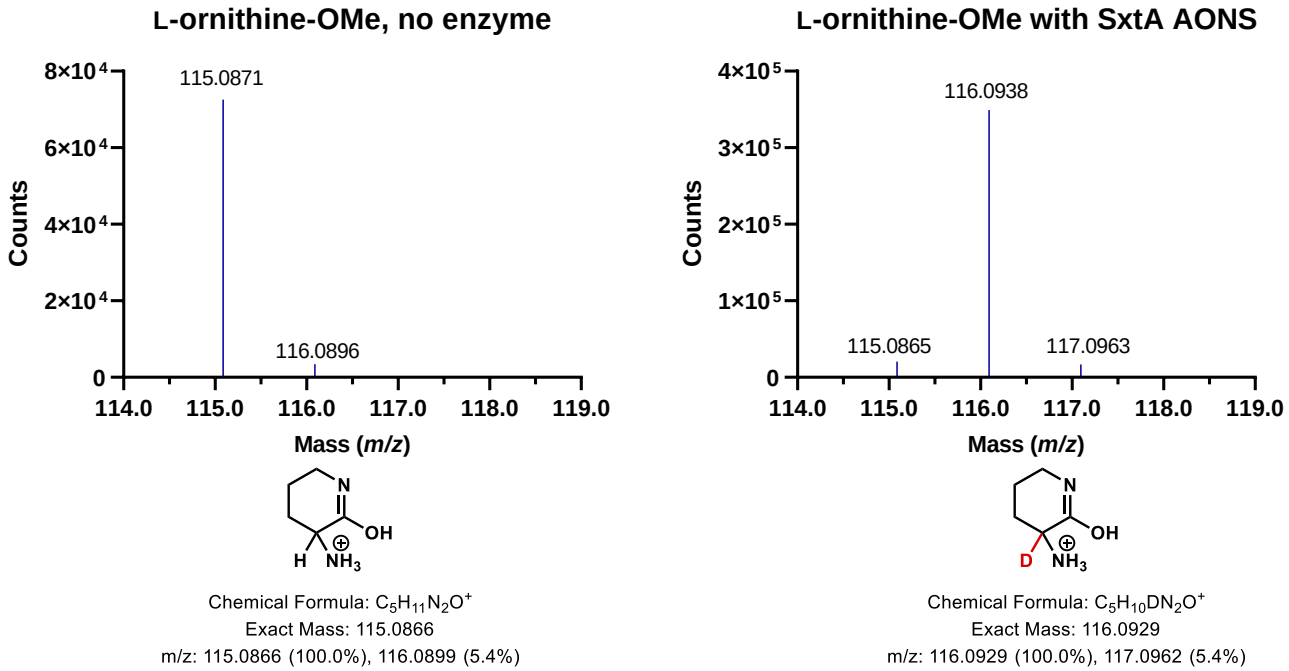
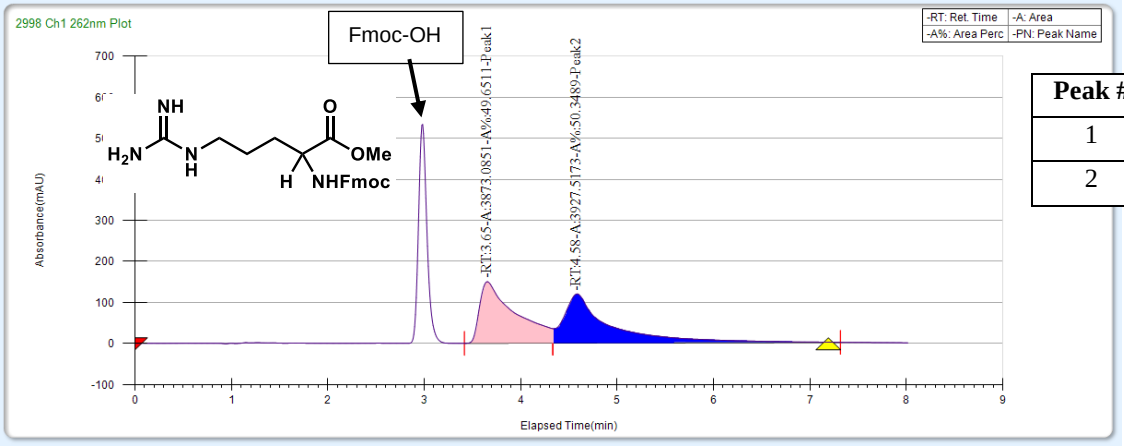


Figure 4.S21. Spectra of ornithine methyl esters.
(A) PDA traces of Fmoc-protected DL-ornithine methyl ester from commercial and synthetic standards, and protected ester after incubation with SxtA AONS (CHIRALCEL OD-H, 40% MeOH with 0.5% formic acid, 3.5 mL/min); (B) representative mass spectra.

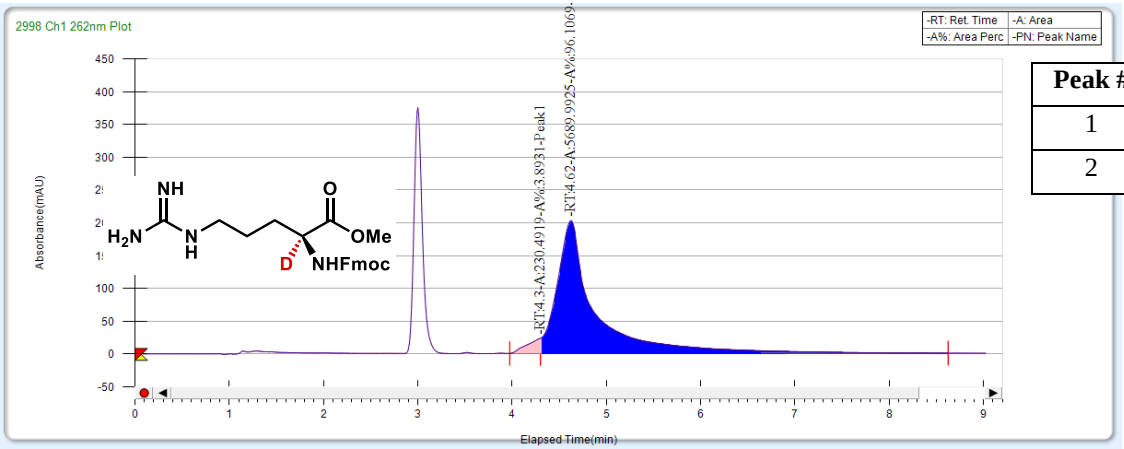
A

Racemic mixture of Arg-OMe



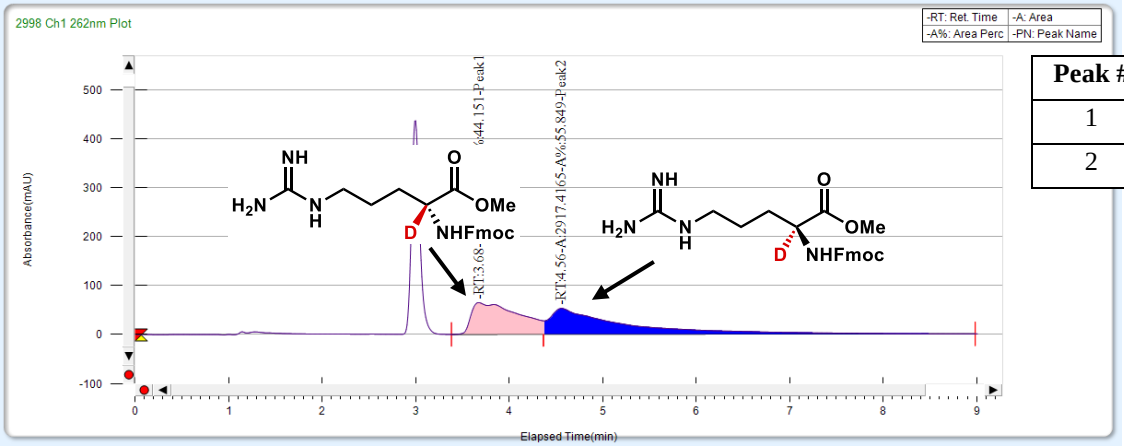
Peak #	Ret. Time	Compound	Area (%)
1	3.7	D-Arg-OMe	49.6511
2	4.6	L-Arg-OMe	50.3489

L-Arg-OMe incubated with SxtA AONS in D_2O



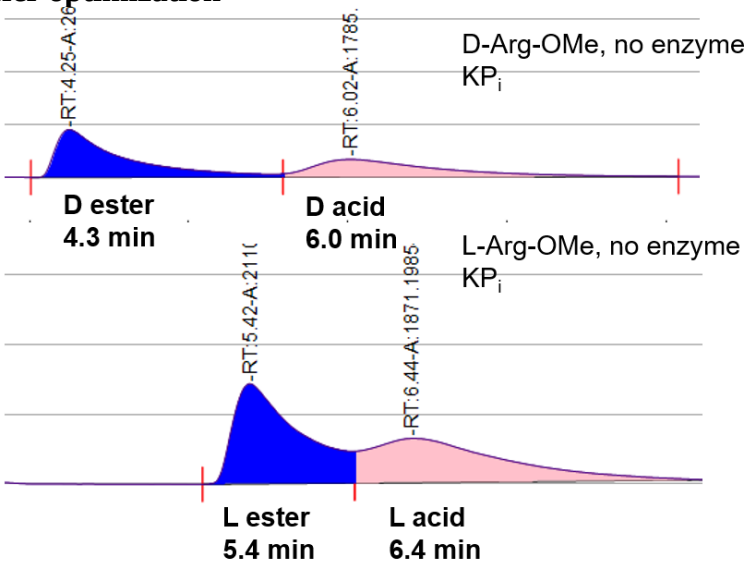
Peak #	Ret. Time	Compound	Area (%)
1	4.3	D-Arg-OMe	3.8931
2	4.6	L-Arg-OMe	96.1069

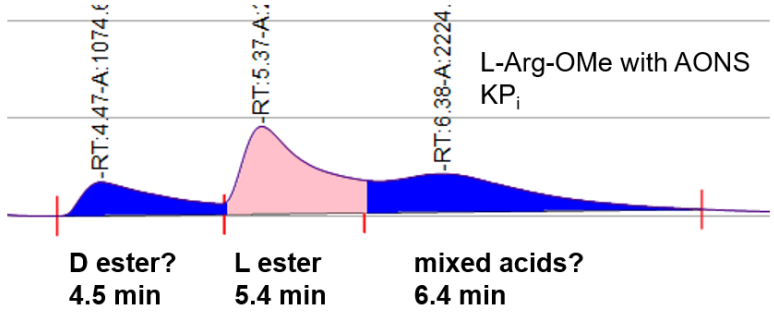
D-Arg-OMe incubated with SxtA AONS in D_2O



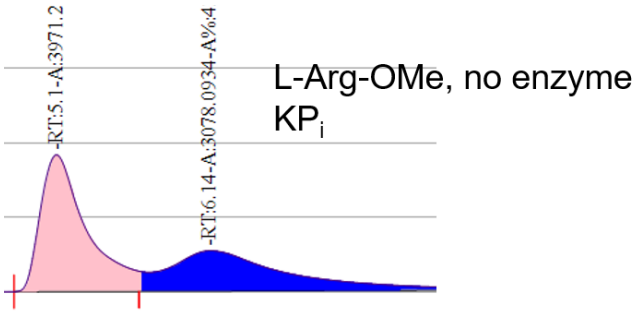
Peak #	Ret. Time	Compound	Area (%)
1	3.7	D-Arg-OMe	44.151
2	4.6	L-Arg-OMe	55.849

Buffer optimization

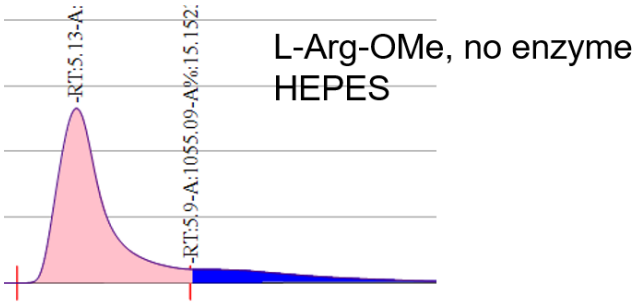




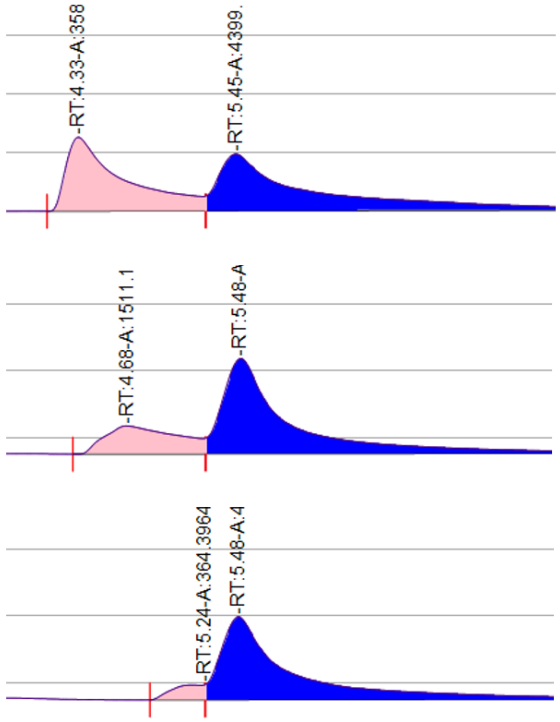
About 50% hydrolysis overnight in phosphate buffer



L-Arg-OMe, no enzyme
KP_i



L-Arg-OMe, no enzyme
HEPES



	D ester	**L ester**
	50% 4.3 min	50% 5.5 min
	25% 4.7 min	75% 5.5 min
	10% 5.2 min	90% 5.5 min

The retention time of D-Arg-OMe shifts later as its relative concentration decreases, while the retention time of L-Arg-OMe remains the same.

B

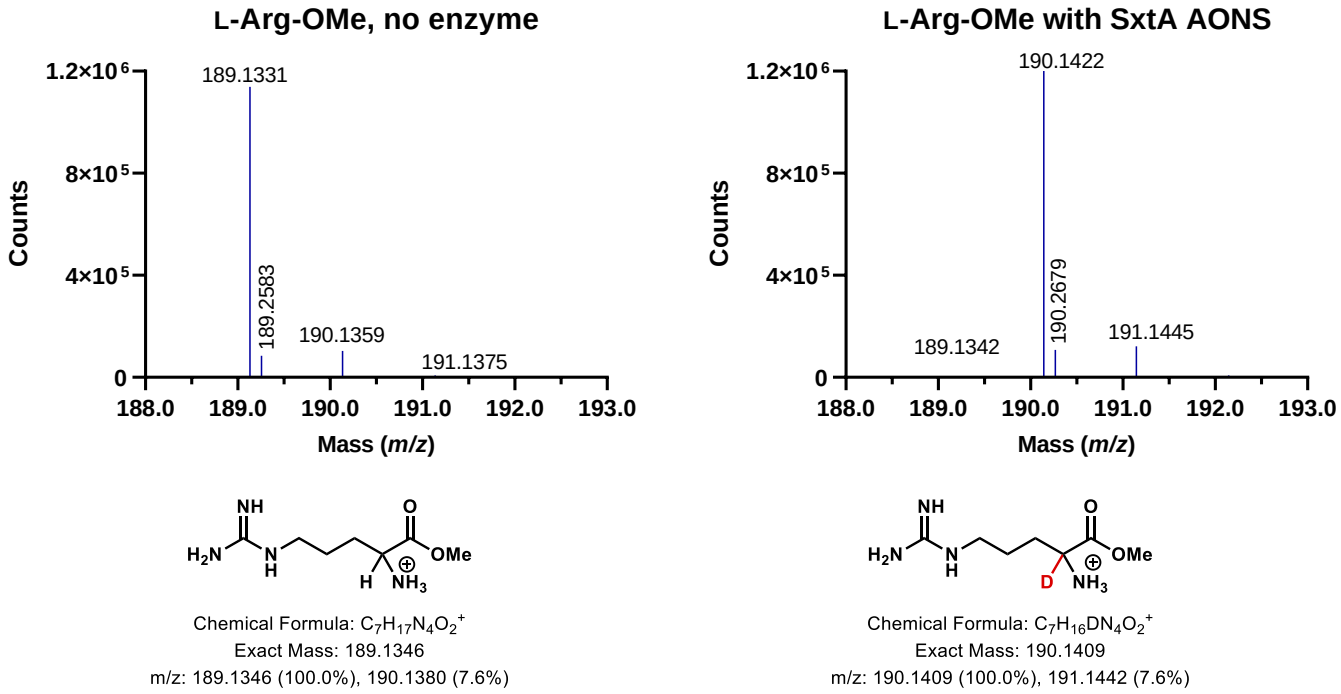


Figure 4.S22. Spectra of arginine methyl esters.
(A) PDA traces of Fmoc-protected DL-arginine methyl ester from commercial standards, and protected ester after incubation with SxtA AONS (CHIRALCEL OD-H, 31% MeOH with 0.5% formic acid, 3.5 mL/min); (B) representative mass spectra.

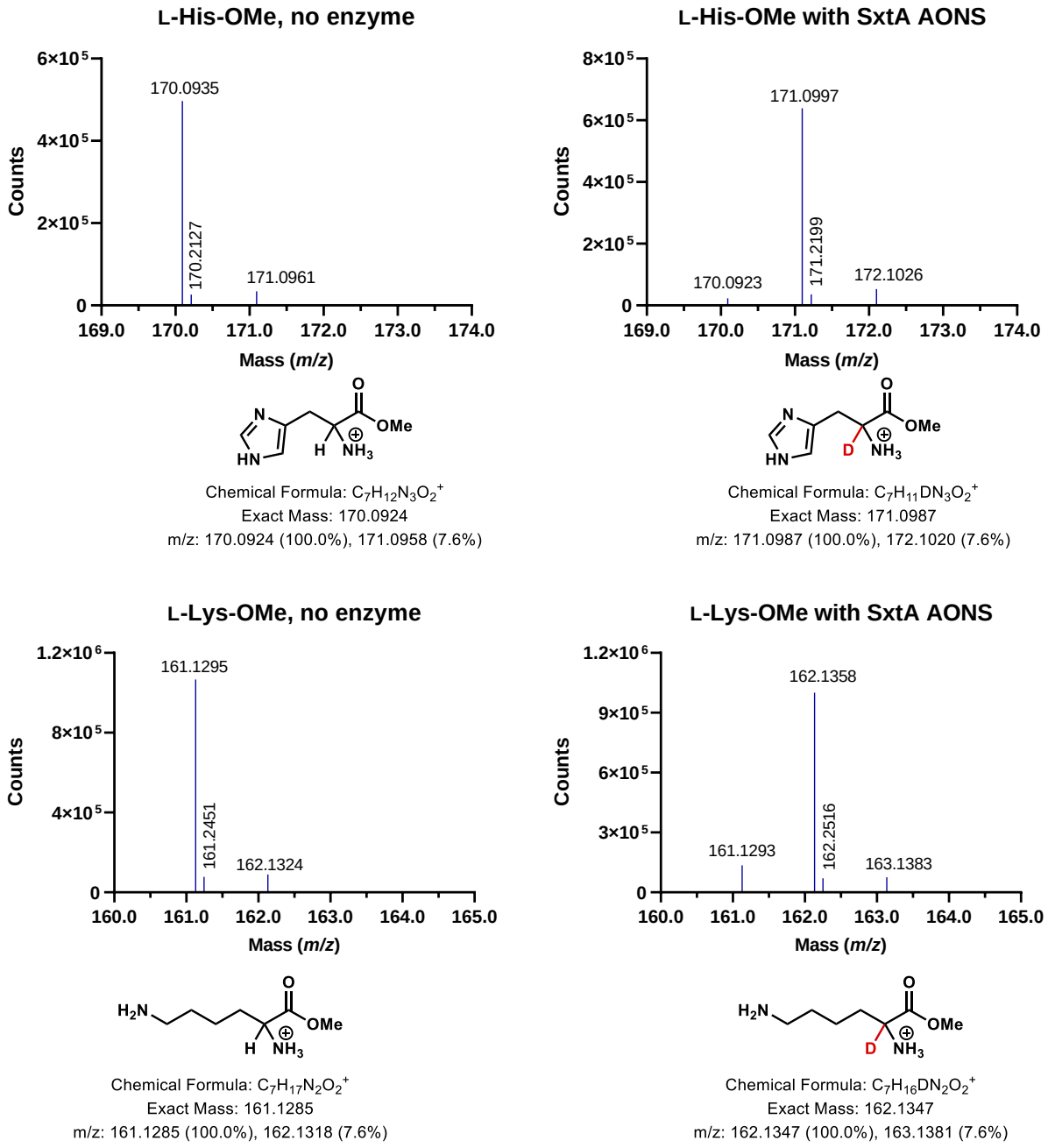


Figure 4.S23. Representative mass spectra of basic α-amino methyl esters (His and Lys) that could not be resolved by SFC

Negative controls (left) and enzymatic reactions with SxtA AONS (right).

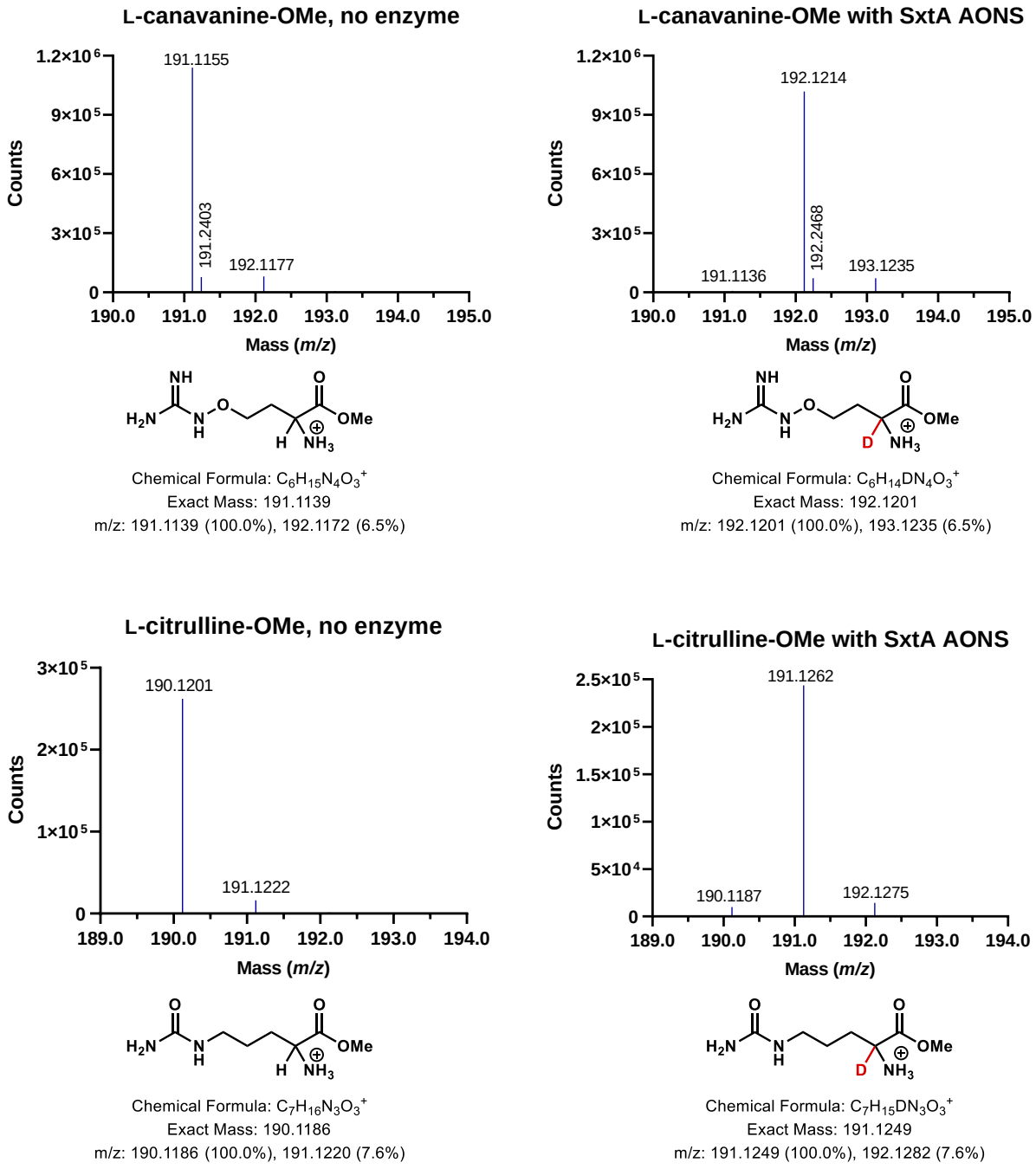


Figure 4.S24. Representative mass spectra of basic α-amino methyl esters (canavanine and citrulline) that could not be resolved by SFC

Negative controls (left) and enzymatic reactions with SxtA AONS (right).

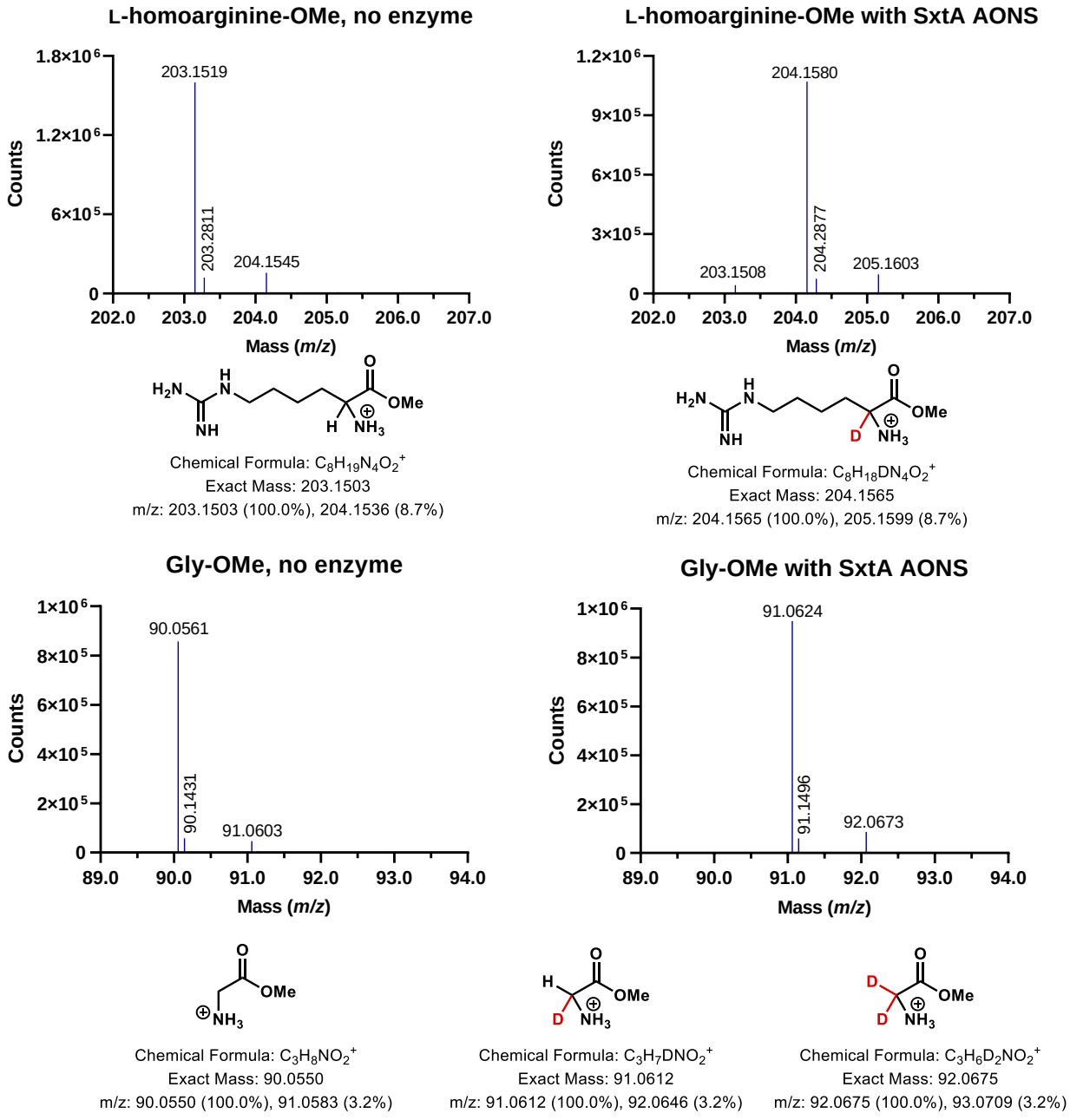


Figure 4.S25. Representative mass spectra of homoarginine methyl esters and Gly-OMe that could not be resolved by SFC.

Negative controls (left) and enzymatic reactions with SxtA AONS (right). About 5% of Gly-OMe showed incorporation of two deuterium atoms.

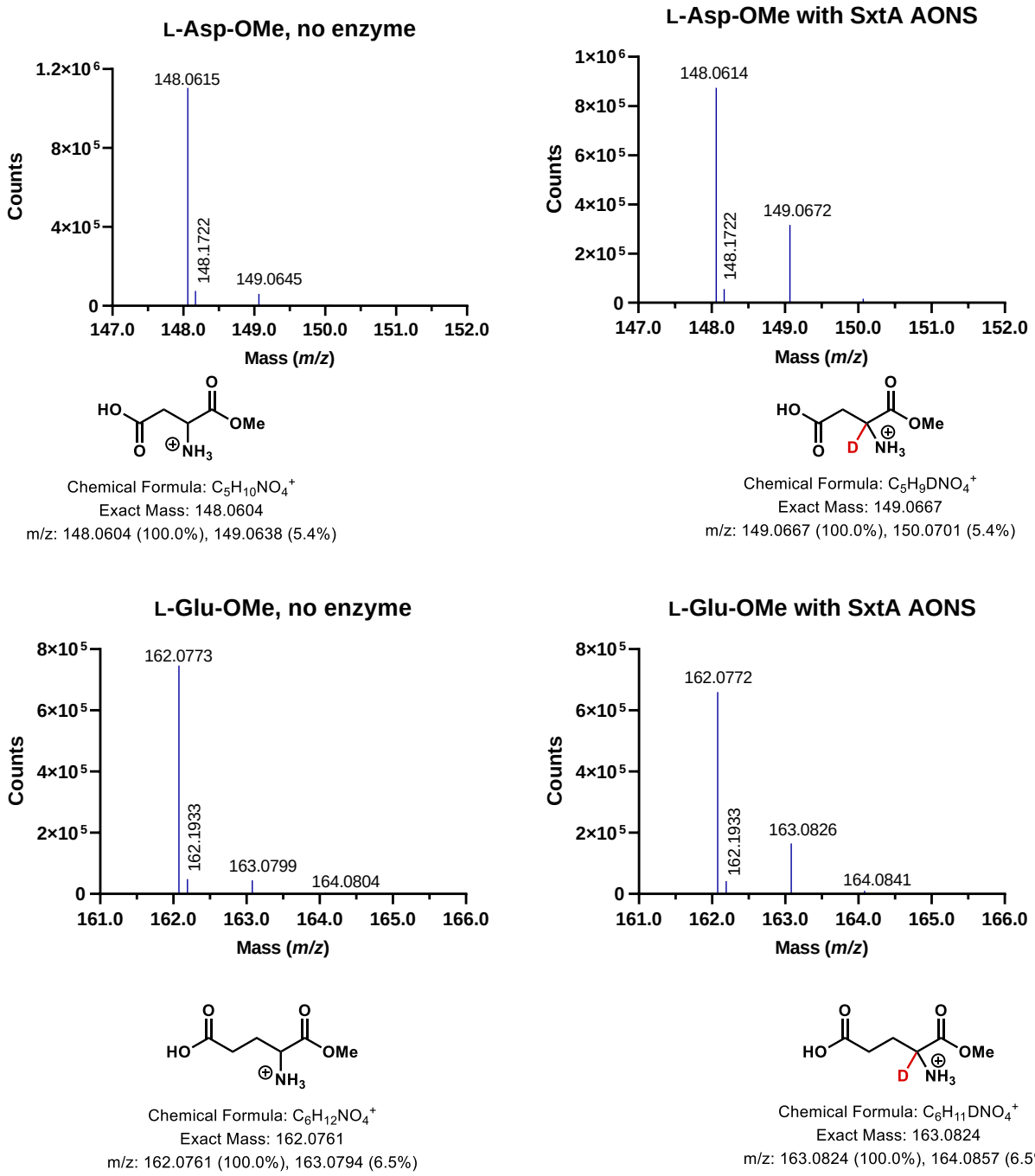


Figure 4.S26. Representative mass spectra of acidic α-amino methyl esters (aspartic and glutamic acids) that could not be resolved by SFC
Negative controls (left) and enzymatic reactions with SxtA AONS (right).

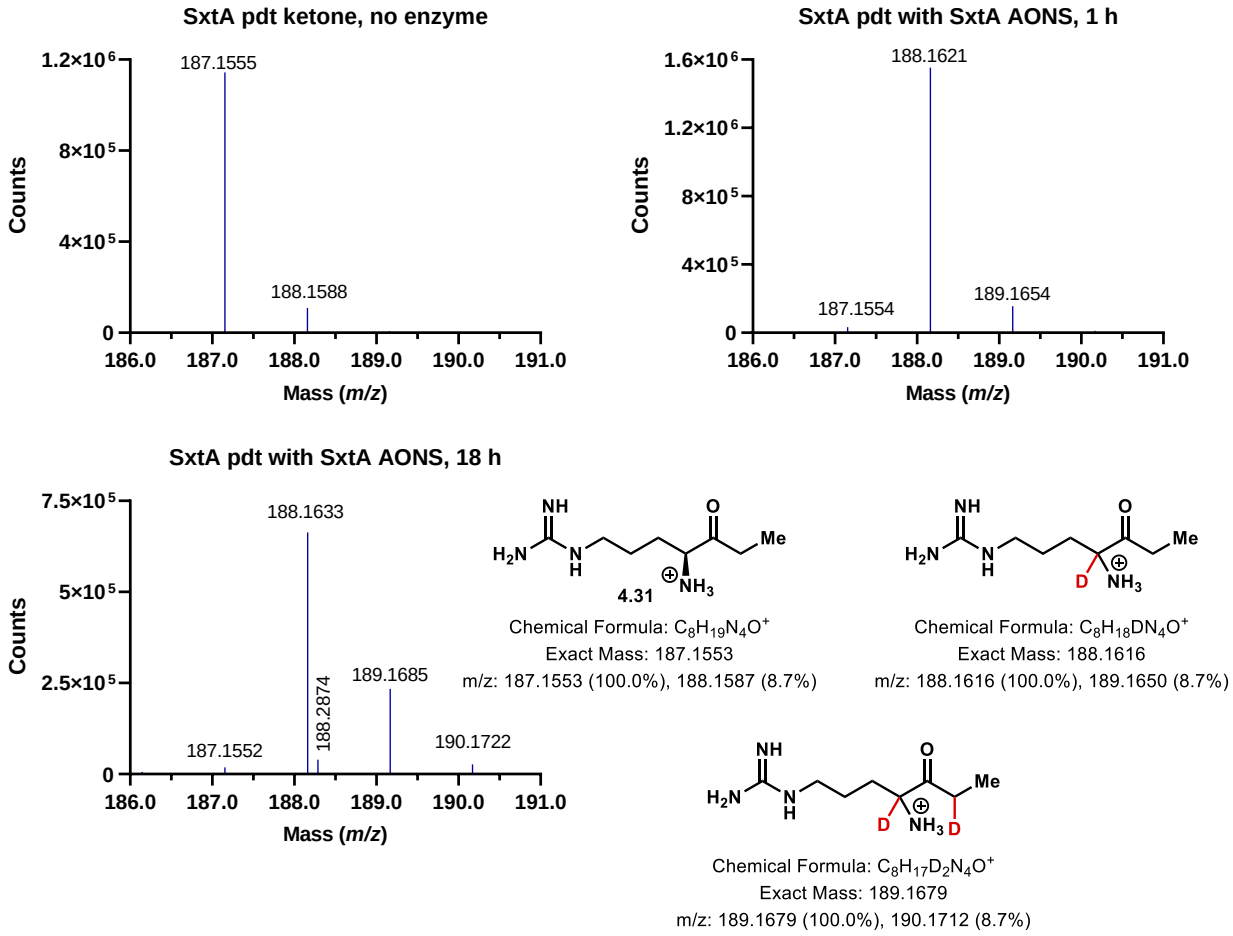


Figure 4.S27. Representative mass spectra of the SxtA product ketone 4.31.
Nonenzymatic negative control and enzymatic reactions with SxtA AONS. At extended timepoints, ketones with more than one deuterium atom incorporated were detected due the additional acidic protons.[5]

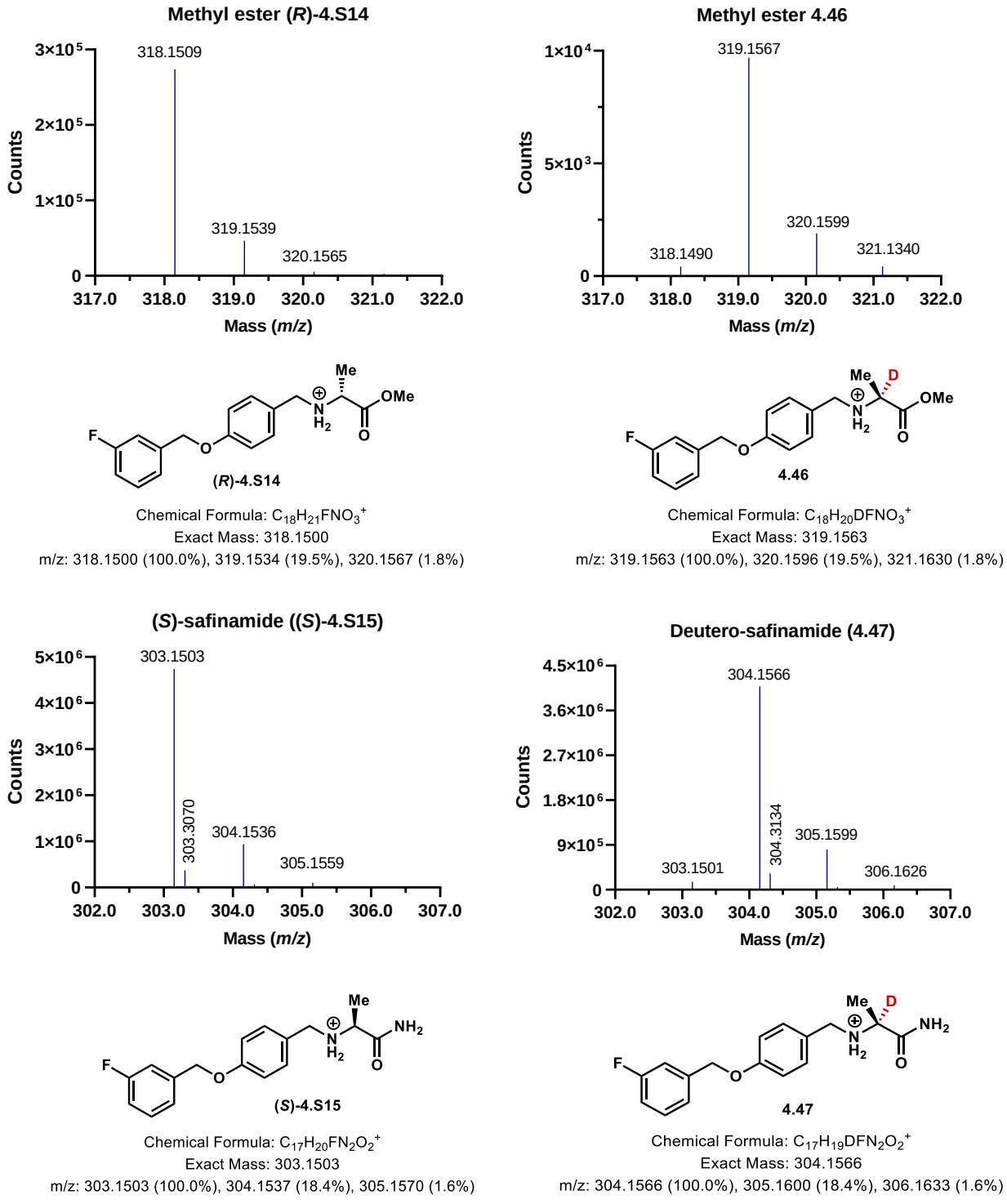


Figure 4.S28. Mass spectra of the safinamide methyl ester precursors 4.S14 and 4.46, and safinamide free bases 4.S15 and 4.47.

V. Proposed mechanisms

The stereochemistry problem we encountered in this Chapter is the related the one observed in Chapter 3 with D-Arg as a condensation substrate. In α-oxoamine synthase (AOS) proteins, the catalytic lysine residue is proposed to stereospecifically mediate the initial deprotonation of the amino acid substrate and the final reprotonation of the alpha position in product ketones. This residue is located on the *si* face of PLP in all published AOS crystal structures, which would result in no deuteration of D-Arg-OMe if it were to proceed through the same mechanism as the L-enantiomer. However, other PLP-dependent enzymes that have both L- and D-amino acid substrates are known: the stereoinverting D-phenylglycine aminotransferase uses alternate binding mode that flips the orientation of the carboxylate group and side chain to place the alpha protons of the D- and L-substrates on the same face as the catalytic lysine,[48] while amino acid racemases have a tyrosine residue on the face opposite to lysine to deprotonate both stereoisomers.[49] It is unclear if SxtA AONS utilizes either of these mechanisms or another alternative, and if the mechanism may change with different substrates.

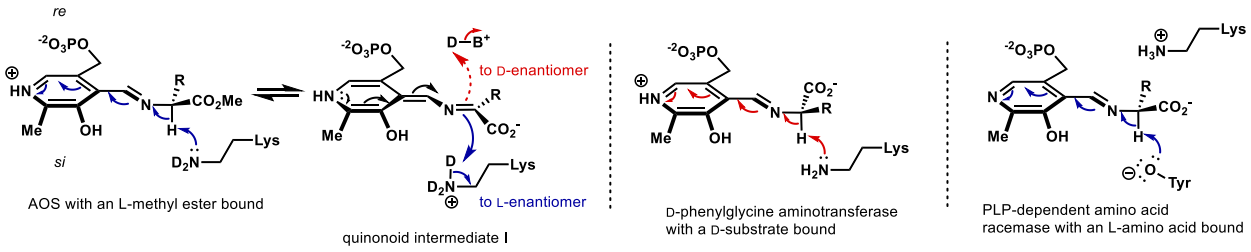


Figure 4.S29. Proposed mechanisms of deprotonation in multiple PLP-dependent enzymes: alternate binding modes or two bases on opposite PLP faces.

Some members of the AOS family are reported to require binding of the thioester partner substrate before deprotonation of the external aldimine I intermediate.[3] The binding of the thioester is proposed rotate the α-proton to be deprotonated by the lysine residue.[42,77] Gratifyingly, the

deuterium incorporation occurred here in the absence of any thioesters that would lead to a mixture of deuterated amino acids and α-amino ketones. Transaminases are reported to deuterate the beta carbons of their substrates,[32] and could generate up to one turnover of undesired α-ketoacid products, which we did not observe with SxtA AONS. Racemases would produce a racemic mixture of products, which is similar to what we saw with D-esters, whereas L-substrates were generally highly stereoretentive.

4.7 References

(1) Pinon, V.; Ravanel, S.; Douce, R.; Alban, C. Biotin Synthesis in Plants. The First Committed Step of the Pathway Is Catalyzed by a Cytosolic 7-Keto-8-Aminopelargonic Acid Synthase. *Plant Physiol.* **2005**, *139*, 1666–1676.

(2) Webster, S. P.; Alexeev, D.; Campopiano, D. J.; Watt, R. M.; Alexeeva, M.; Sawyer, L.; Baxter, R. L. Mechanism of 8-Amino-7-Oxononanoate Synthase: Spectroscopic, Kinetic, and Crystallographic Studies. *Biochemistry* **2000**, *39*, 516–528.

(3) Ikushiro, H.; Fujii, S.; Shiraiwa, Y.; Hayashi, H. Acceleration of the Substrate Cα Deprotonation by an Analogue of the Second Substrate Palmitoyl-CoA in Serine Palmitoyltransferase. *J. Biol. Chem.* **2008**, *283*, 7542–7553.

(4) Kaufholz, A.; Hunter, G. A.; Ferreira, G. C.; Lendrihas, T.; Hering, V.; Layer, G.; Jahn, M.; Jahn, D. Aminolaevulinic Acid Synthase of Rhodobacter Capsulatus : High-Resolution Kinetic Investigation of the Structural Basis for Substrate Binding and Catalysis. *Biochem. J.* **2013**, *451*, 205–216.

(5) Ploux, O.; Marquet, A. Mechanistic Studies on the 8-Amino-7-Oxopelargonate Synthase, a Pyridoxal-5'-Phosphate-Dependent Enzyme Involved in Biotin Biosynthesis. *Eur. J. Biochem.* **1996**, *236*, 301–308.

(6) Nandi, D. L. Studies on Delta-Aminolevulinic Acid Synthase of Rhodopseudomonas Sphaeroides. *J. Biol. Chem.* **1978**, *253*, 8872–8877.

(7) Stojanovski, B. M.; Hunter, G. A.; Jahn, M.; Jahn, D.; Ferreira, G. C. Unstable Reaction Intermediates and Hysteresis during the Catalytic Cycle of 5-Aminolevulinate Synthase: Implications from Using Pseudo and Alternate Substrates and a Promiscuous Enzyme Variant. *J. Biol. Chem.* **2014**, *289*, 22915–22925.

(8) Zaman, Z.; Jordan, P. M.; Akhtar, M. Mechanism and Stereochemistry of the 5 Aminolaevulinate Synthetase Reaction. *Biochem. J.* **1973**, *135*, 257–263.

(9) Alexeev, D.; Baxter, R. L.; Campopiano, D. J.; Kerbarh, O.; Sawyer, L.; Tomczyk, N.; Webster, S. P. Suicide Inhibition of a -Oxamine Synthases : Structures of the Covalent Adducts of 8-Amino-7-Oxononanoate Synthase with Trifluoroalanine. *Org. Biomol. Chem.* **2006**, *4*, 1209–1212.

(10) Harbeson, S. L.; Tung, R. D. Deuterium in Drug Discovery and Development. *Annu. Rep. Med. Chem.* **2011**, *46*, 403–417.

(11) Westheimer, F. H. The Magnitude of the Primary Kinetic Isotope Effect for Compounds of Hydrogen and Deuterium. *Chem. Rev.* **1961**, *61*, 265–273.

(12) Pony Yu, R.; Hesk, D.; Rivera, N.; Pelczer, I.; Chirik, P. J. Iron-Catalysed Tritiation of Pharmaceuticals. *Nature* **2016**, *529*, 195–199.

(13) Loh, Y. Y.; Nagao, K.; Hoover, A. J.; Hesk, D.; Rivera, N. R.; Colletti, S. L.; Davies, I. W.; MacMillan, D. W. C. Photoredox-Catalyzed Deuteration and Tritiation of Pharmaceutical Compounds. *Science* **2017**, *358*, 1182–1187.

(14) Timmins, G. S. Deuterated Drugs; Updates and Obviousness Analysis. *Expert Opin. Ther. Pat.* **2017**, *27*, 1353–1361.

(15) Darland, G. K.; Hajdu, R.; Kropp, H.; Kahan, F. M.; Walker, R. W.; Vandenheuvel, W. J. A. Oxidative and Defluorinative Metabolism of Fludalanine, 2-2H-3-Fluoro-D-Alanine. *Drug Metab. Dispos.* **1986**, *14*, 668–673.

(16) Reider, P. J.; Eichen Conn, R. S.; Davis, P.; Grenda, V. J.; Zambito, A. J.; Grabowski, E. J. J. Synthesis of (R)-Serine-2-d and Its Conversion to the Broad Spectrum Antibiotic Fludalanine. *J. Org. Chem.* **1987**, *52*, 3326–3334.

(17) Dang, T.-Y.; Cheung, Y.-F.; Walsh, C. Reactions of β-Fluoroalanine and β-Bromoalanine with D-Amino Acid Oxidase. *Biochem. Biophys. Res. Commun.* **1976**, *72*, 960–968.
(18) Dean, M.; Sung, V. W. Review of Deutetrabenazine: A Novel Treatment for Chorea Associated with Huntington's Disease. *Drug Des. Devel. Ther.* **2018**, *12*, 313–319.
(19) de la Torre, B. G.; Albericio, F. The Pharmaceutical Industry in 2017. An Analysis of FDA Drug Approvals from the Perspective of Molecules. *Molecules* **2018**, *23*, 533.
(20) Sheppard, D.; Li, D. W.; Brüschweiler, R.; Tugarinov, V. Deuterium Spin Probes of Backbone Order in Proteins: 2H NMR Relaxation Study of Deuterated Carbon α Sites. *J. Am. Chem. Soc.* **2009**, *131*, 15853–15865.
(21) Bornø, A.; Van Hall, G. Quantitative Amino Acid Profiling and Stable Isotopically Labeled Amino Acid Tracer Enrichment Used for in Vivo Human Systemic and Tissue Kinetics Measurements. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **2014**, *951–952*, 69–77.
(22) Bode, H. B.; Brachmann, A. O.; Jadhav, K. B.; Seyfarth, L.; Dauth, C.; Fuchs, S. W.; Kaiser, M.; Waterfield, N. R.; Sack, H.; Heinemann, S. H.; Arndt, H.-D. Structure Elucidation and Activity of Kolossin A , the D-/L-Pentadecapeptide Product of a Giant Nonribosomal Peptide Synthetase. *Angew. Chem. Int. Ed.* **2015**, *54*, 10352–10355.
(23) Seebach, D.; Dziadulewiczl, E.; Behrendt, L.; Cantoreggi, S.; Fitzi, R. Synthesis of Nonproteinogenic (R)- or (S)-Amino Acids Analogues of Phenylalanine, Isotopically Labeled and Cyclic Amino Acids from Tert-Butyl 2-(Tert-Butyl)-3-Methyl-4-Oxo-1-Imidazolinecarboxylate (Boc-BMI). *Liebigs Ann. der Chemie* **1989**, 1215–1232.
(24) Hoppe, D.; Ludger, B. Selective Mono- and Dialkylation of N-[Bis(Alkylthio)Methylene]Glycine Ethyl Ester for Synthesis of Higher and α-Branched α-Amino Acids. *Liebigs Ann. der Chemie* **1979**, 2066–2075.
(25) Elemes, Y.; Ragnarsson, U. Synthesis of Enantiopure Alpha-Deuteriated Boc-L-Amino Acids. *J. Chem. Soc. Perkin Trans. 1* **1996**, 537–540.
(26) Takeda, R.; Abe, H.; Shibata, N.; Moriwaki, H.; Izawa, K.; Soloshonok, V. A. Asymmetric Synthesis of Alpha-Deuterated Alpha-Amino Acids. *Org. Biomol. Chem.* **2017**, *15*, 6978–6983.
(27) Chatterjee, B.; Krishnakumar, V.; Gunanathan, C. Selective Alpha-Deuteration of Amines and Amino Acids Using D2O. *Org. Lett.* **2016**, *18*, 5892–5895.
(28) Truppo, M. D. Biocatalysis in the Pharmaceutical Industry : The Need for Speed. *ACS Med. Chem. Lett.* **2017**, *8*, 476–480.
(29) $15.80 and $1,165 per Liter of D2 and D2O, Respectively on Sigma-Aldrich in April 2020.
(30) Dunathan, H. C. Conformation and Reaction Specificity in Pyridoxal Phosphate Enzymes. *Proc. Natl. Acad. Sci.* **1966**, *55*, 712–716.
(31) Toney, M. D. Controlling Reaction Specificity in Pyridoxal Phosphate Enzymes. *Biochim. Biophys. Acta - Proteins Proteomics* **2011**, *1814*, 1407–1418.
(32) Babu, U. M.; Johnston, R. B. D2O-Alanine Exchange Reactions Catalyzed by Alanine Racemase and Glutamic Pyruvic Transaminase. *Biochem. Biophys. Res. Commun.* **1974**, *58*, 460–466.
(33) Lim, Y.-H.; Yoshimura, T.; Soda, K.; Esaki, N. Stereospecific Labeling at Alpha-Position of Phenylalanine Phenylglycine with Amino Acid Racemase. *J. Ferment. Bioeng.* **1998**, *86*, 400–402.
(34) Bette, S.; Shpiner, D. S.; Singer, C.; Moore, H. Safinamide in the Management of Patients with Parkinson's Disease Not Stabilized on Levodopa: A Review of the Current Clinical Evidence. *Ther. Clin. Risk Manag.* **2018**, *14*, 1737–1745.

(35) Salvati, P.; Maj, R.; Caccia, C.; Cervini, M. A.; Fornaretto, M. G.; Lamberti, E.; Pevarello, P.; Skeen, G. A.; White, H. S.; Wolf, H. H.; Faravelli, L.; Mazzanti, M.; Mancinelli, E.; Varasi, M.; Fariello, R. G. Biochemical and Electrophysiological Studies on the Mechanism of Action of PNU-151774E, a Novel Antiepileptic Compound. *J. Pharmacol. Exp. Ther.* **1999**, *288*, 1151–1159.
(36) Gómez-Gallego, M.; Sierra, M. A. Kinetic Isotope Effects in the Study of Organometallic Reaction Mechanisms. *Chem. Rev.* **2011**, *111*, 4857–4963.
(37) Gong, J.; Hunter, G. A.; Ferreira, G. C. Aspartate-279 in Aminolevulinate Synthase Affects Enzyme Catalysis through Enhancing the Function of the Pyridoxal 5′-Phosphate Cofactor. *Biochemistry* **1998**, *2960*, 3509–3517.
(38) Kerbarh, O.; Campopiano, D. J.; Baxter, R. L. Mechanism of Alpha-Oxoamine Synthases: Identification of the Intermediate Claisen Product in the 8-Amino-7-Oxononanoate Synthase Reaction. *Chem. Commun.* **2006**, *18*, 60–62.
(39) Cooper, A. J. L. Proton Magnetic Resonance Studies of Glutamate-Alanine Deuterium Exchange. *J. Biol. Chem.* **1976**, *251*, 1088–1097.
(40) Golichowski, A.; Harruff, R. C.; Jenkins, W. T. The Effects of PH on the Rates of Isotope Exchange Alanine Aminotransferase. *Arch. Biochem. Biophys.* **1977**, *178*, 459–467.
(41) Hunter, G. A.; Ferreira, G. C. Lysine-313 of 5-Aminolevulinate Synthase Acts as a General Base during Formation of the Quinonoid Reaction Intermediates. *Biochemistry* **1999**, *38*, 3711–3718.
(42) Raman, M. C. C.; Johnson, K. A.; Yard, B. A.; Lowther, J.; Carter, L. G.; Naismith, J. H.; Campopiano, D. J. The External Aldimine Form of Serine Palmitoyltransferase. *J. Biol. Chem.* **2009**, *284*, 17328–17339.
(43) Alexeev, D.; Alexeeva, M.; Baxter, R. L.; Campopiano, D. J.; Webster, S. P.; Sawyer, L. The Crystal Structure of 8-Amino-7-Oxononanoate Synthase: A Bacterial PLP-Dependent, Acyl-CoA-Condensing Enzyme. *J. Mol. Biol.* **1998**, *284*, 401–419.
(44) Schmidt, A.; Sivaraman, J.; Li, Y.; Larocque, R.; Barbosa, J. A. R. G.; Smith, C.; Matte, A.; Schrag, J. D.; Cygler, M. Three-Dimensional Structure of 2-Amino-3-Ketobutyrate CoA Ligase from *Escherichia Coli* Complexed with a PLP−Substrate Intermediate: Inferred Reaction Mechanism. *Biochemistry* **2001**, *40*, 5151–5160.
(45) Astner, I.; Schulze, J. O.; van den Heuvel, J.; Jahn, D.; Schubert, W.-D.; Heinz, D. W. Crystal Structure of 5-Aminolevulinate Synthase, the First Enzyme of Heme Biosynthesis, and Its Link to XLSA in Humans. *EMBO J.* **2005**, *24*, 3166–3177.
(46) Yard, B. A.; Carter, L. G.; Johnson, K. A.; Overton, I. M.; Dorward, M.; Liu, H.; McMahon, S. A.; Oke, M.; Puech, D.; Barton, G. J.; Naismith, J. H.; Campopiano, D. J. The Structure of Serine Palmitoyltransferase; Gateway to Sphingolipid Biosynthesis. *J. Mol. Biol.* **2007**, *370*, 870–886.
(47) Jahan, N.; Potter, J. A.; Sheikh, M. A.; Botting, C. H.; Shirran, S. L.; Westwood, N. J.; Taylor, G. L. Insights into the Biosynthesis of the Vibrio Cholerae Major Autoinducer CAI-1 from the Crystal Structure of the PLP-Dependent Enzyme CqsA. *J. Mol. Biol.* **2009**, *392*, 763–773.
(48) Walton, C. J. W.; Thiebaut, F.; Brunzelle, J. S.; Couture, J.-F.; Chica, R. A. Structural Determinants of the Stereoinverting Activity of Pseudomonas Stutzeri D-Phenylglycine Aminotransferase. *Biochemistry* **2018**, *57*, 5437–5446.
(49) Spies, M. A.; Toney, M. D. Multiple Hydrogen Kinetic Isotope Effects for Enzymes Catalyzing Exchange with Solvent: Application to Alanine Racemase. *Biochemistry* **2003**,

42, 5099–5107.
(50) Richard, J. P.; Amyes, T. L.; Crugeiras, J.; Rios, A. Pyridoxal 5′-Phosphate: Electrophilic Catalyst Extraordinaire. *Curr. Opin. Chem. Biol.* **2009**, *13*, 475–483.
(51) Rios, A.; Amyes, T. L.; Richard, J. P. Formation and Stability of Organic Zwitterions in Aqueous Solution: Enolates of the Amino Acid Glycine and Its Derivatives. *J. Am. Chem. Soc.* **2000**, *122*, 9373–9385.
(52) Pugnière, M.; Commeyras, A.; Previero, A. Racemization of Amino Acid Esters Catalysed by Pyridoxal 5′ Phosphate as a Step in the Production of L-Amino Acids. *Biotechnol. Lett.* **1983**, *5*, 447–452.
(53) Zabinski, R. F.; Toney, M. D. Metal Ion Inhibition of Nonenzymatic Pyridoxal Phosphate Catalyzed Decarboxylation and Transamination. *J. Am. Chem. Soc.* **2001**, *123*, 193–198.
(54) Cao, X.; Chen, H.; Du, Q.; Shen, X. High-Purity Safinamide Preparing Method. CN105061245A, 2015.
(55) Barbanti, E.; Faravelli, L.; Salvati, P.; Canevotti, R.; Ponzini, F. Process for the Production of 2-[4-(3- or 2-Fluorobenzyloxy)Benzylamino]Propanamides with High Purity Degree. WO2009074478 A1, 2009.
(56) Brückner, H.; Sorsche, B.; Esna-Ashari, A.; Jöster, R. Chiral Ligand-Exchange Chromatography of Amino Acid Derivatives. In *Amino Acids*; Springer Netherlands: Dordrecht, 1990; pp 152–158.
(57) El-Faham, A.; Albericio, F. Peptide Coupling Reagents, More than a Letter Soup. *Chem. Rev.* **2011**, *111*, 6557–6602.
(58) Kazemi, M.; Shiri, L. Thioesters Synthesis : Recent Adventures in the Esterification of Thiols. *J. Sulfur Chem.* **2015**, *35*, 613–623.
(59) Jain, M.; Barthwal, M. K.; Haq, W.; Katti, S. B.; Dikshit, M. Synthesis and Pharmacological Evaluation of Novel Arginine Analogs as Potential Inhibitors of Acetylcholine-Induced Relaxation in Rat Thoracic Aortic Rings. *Chem. Biol. Drug Des.* **2012**, *79*, 459–469.
(60) Boddey, J.; Cowman, A.; Czabotar, P.; Hodder, T.; Sleebs, B. Structure of Plasmepsin V in Complex with an Inhibitor and Uses Thereof. WO 2016/197190 A1, 2016.
(61) Guo, L.; Wang, C.; Zhao, W.; Li, H.; Sun, W.; Shen, Z. Copolymerization of CO2 and Cyclohexene Oxide Using a Lysine-Based (Salen)Cr(III)Cl Catalyst. *Dalt. Trans.* **2009**, *2*, 5406–5410.
(62) Ambrosi, H.; Hartmann, V.; Pistorius, D.; Reissbrodt, R.; Trowitzsch-Kienast, W. Myxochelins B, C, D, E and F: A New Structural Principle for Powerful Siderophores Imitating Nature. *European J. Org. Chem.* **1998**, *1998*, 541–551.
(63) Pei, Z.; Toone, E. J. Ethynylbenzene Derivatives. WO 2012/031298 A2, 2012.
(64) Lloyd-Williams, P.; Monerris, P.; Gonzalez, I.; Jou, G.; Giralt, E. Synthesis of D-Alloisoleucine from L-Isoleucine and from (S)-2-Methylbutan-1-Ol. Synthesis of Isostatine. *J. Chem. Soc. Perkins Trans. 1* **1994**, 1969–1974.
(65) Mojmir, S.; Li, A.; Liu, Y.; Feng, Q.; Bartha, R.; Hudson, R. Preliminary Evaluation of PARACREST MRI Agents for the Detection of Nitric Oxide Synthase. *Can. J. Chem.* **2016**, *94*, 715–722.
(66) Kuttan, A.; Nowshudin, S.; Rao, M. N. A. Ceric Ammonium Nitrate (CAN) Mediated Esterification of N-Boc Amino Acids Allows Either Retention or Removal of the N-Boc Group. *Tetrahedron Lett.* **2004**, *45*, 2663–2665.
(67) Howes, C.; Alcock, N. W.; Golding, B. T.; McCabe, R. W. Cyclo-(L-Asparagyl-L-Asparagyl) [(3S,6S)-3,6-Bis(Carbamoylmethyl)-Piperazine-2,5-Dione]: Preparation and

Crystal Structure. *J. Chem. Soc. Perkins Trans. 1* **1983**, 2287–2291.
(68) Castonguay, R.; Lherbet, C.; Keillor, J. W. Mapping of the Active Site of Rat Kidney Gamma-Glutamyl Transpeptidase Using Activated Esters and Their Amide Derivatives. *Bioorganic Med. Chem.* **2002**, *10*, 4185–4191.
(69) Glutamic acid α-methyl ester http://www.chemimpex.com/l-glutamic-acid-alpha-methyl-ester.
(70) Tsuchiya, S.; Cho, Y.; Konoki, K.; Nagasawa, K.; Oshima, Y.; Yotsu-Yamashita, M. Synthesis and Identification of Proposed Biosynthetic Intermediates of Saxitoxin in the Cyanobacterium Anabaena Circinalis (TA04) and the Dinoflagellate Alexandrium Tamarense (Axat-2). *Org. Biomol. Chem.* **2014**, *12*, 3016–3020.
(71) Johnson, M. R.; Thelin, W. R.; Angust, J. R. A. Dithiol Mucolytic Agents. US2015056305, 2015.
(72) O'Reilly, E.; Balducci, D.; Paradisi, F. A Stereoselective Synthesis of α-Deuterium-Labelled (S)-α-Amino Acids. *Amino Acids* **2010**, *39*, 849–858.
(73) Maria, R.; Di, C.; Simone, A. De; Andrisano, V.; Bisignano, P.; Bisi, A.; Gobbi, S.; Rampa, A.; Fato, R.; Bergamini, C.; Perez, D. I.; Martinez, A.; Bottegoni, G.; Cavalli, A.; Belluti, F. Versatility of the Curcumin Scaffold: Discovery of Potent and Balanced Dual BACE-1 and GSK-3 β Inhibitors. *J. Med. Chem.* **2016**, *59*, 531–544.
(74) Eschenfeldt, W. H.; Lucy, S.; Millard, C. S.; Joachimiak, A.; Mark, I. D. A Family of LIC Vectors for High-Throughput Cloning and Purification of Proteins. In *Methods in Molecular Biology*; Doyle, S. A., Ed.; Humana Press: Totowa, NJ, 2009; Vol. 498, pp 105–115.
(75) Hill, M. P.; Carroll, E. C.; Vang, M. C.; Addington, T. A.; Toney, M. D.; Larsen, D. S. Light-Enhanced Catalysis by Pyridoxal Phosphate-Dependent Aspartate Aminotransferase. *J. Am. Chem. Soc.* **2010**, *132*, 16953–16961.
(76) Zhou, W.; Zhang, X. Y.; Duan, G. L. Liquid-Chromatography Quantitative Analysis of 20 Amino Acids after Derivatization with FMOC-CI and Its Application to Different Origin Radix Isatidis. *J. Chinese Chem. Soc. Taipei* **2011**, *58*, 509–515.
(77) Shiraiwa, Y.; Ikushiro, H.; Hayashi, H. Multifunctional Role of His159 in the Catalytic Reaction of Serine Palmitoyltransferase. *J. Biol. Chem.* **2009**, *284*, 15487–15495.

Chapter 5: Structural and Spectroscopic Characterization of SxtA

Sarah Ackenhusen, Dr. Meredith Skiba and Yongtong Shero Lao contributed to experiments in this Chapter.

Summary

After reporting the native function of the SxtA, we have worked to develop the individual partner enzyme domains for productive enzymatic reactions on non-native substrates. Finer mechanistic details of the SxtA MT and AONS reactions that may resolve previously observed limitations in reactivity have yet to be elucidated. Here, we discuss structural and spectroscopic studies of these two catalytic domains. SxtA MT was found to be a metal-dependent methyltransferase but employs Mn^{2+} for monomethylation of malonyl-ACP compared to dimethylation of the same substrate by other GNAT-associated MTs, AprA and GphF MT_L. With a homology structure of SxtA MT, we sought to identify the origin of these activity differences. We have not observed SxtA MT-catalyzed dimethylation. Additionally, we conducted spectroscopic studies on SxtA AONS to identify potential rate-determining or unproductive steps that limit the enzyme's activity. Spectroscopic changes in the presence of the native substrate, product and some compatible thiols were observed, and key intermediate species were detected. The residues for productive protein-small molecule interactions and conformational changes have not been identified yet. These results will be used to improve SxtA AONS' biocatalytic potential in the transformation of α-amino acids to α-amino ketones, α-deuterated amino acids and tetra-substituted amino acid derivatives.

5.1 Introduction

After probing the reactivity patterns of SxtA domains Chapters 3 and 4, we are also interested in the structural features of its domains and inter-domain protein-protein interactions that dictate its functions (Figure 5.1). Further understanding of the mechanisms of SxtA domains could contribute to future improvement of the individual domains or the entire SxtA module to make more robust and promiscuous biocatalysts. Of particular interest are the SxtA MT and AONS domains, which belong to subsets of larger protein classes, methyltransferases and PLP-dependent enzymes. The general properties of these protein families have been reviewed elsewhere,[1,2] but much less is known about the structural and mechanistic details of both GNAT-associated MTs (MT_Ls) like SxtA MT and α-oxoamine synthases (AOS).

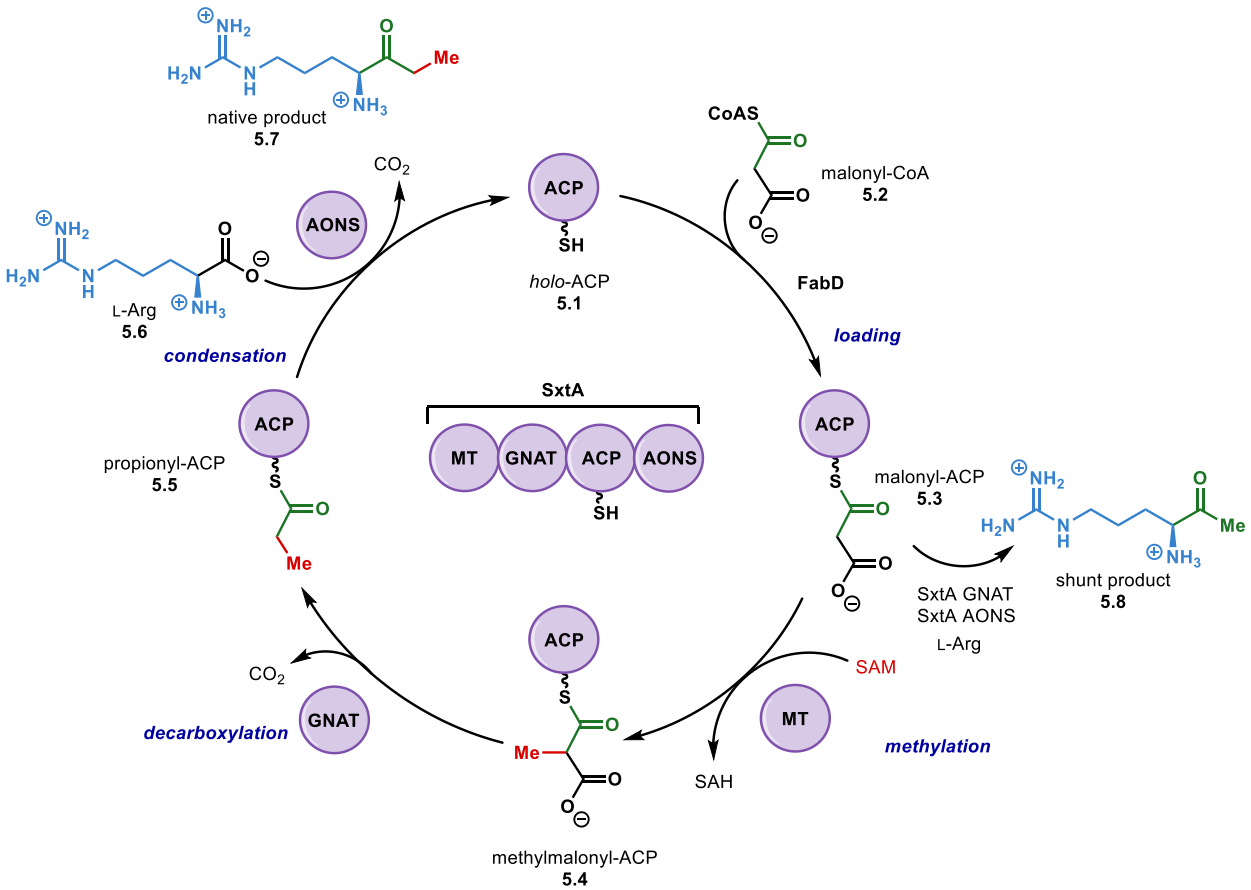


Figure 5.1. The catalytic cycle of the SxtA module leading to formation of the native and shunt ketone products.

Many of the experiments discussed in this Chapter were designed with the assistance and input of Prof. Janet L. Smith's group at the University of Michigan. Smith lab members Dr. Meredith Skiba and later Yongtong Shero Lao had already been studying other GNAT-associated MTs for years when we began investigating SxtA in late 2016 and agreed to collaborate on structural studies of SxtA MT and AONS domains.

Chapter 5 presents our efforts to understand finer mechanistic details of SxtA MT and AONS. For SxtA MT, this was based on a sequence alignment and a homology model. We do not have additional structural data of SxtA AONS, so we summarize our spectroscopic studies to detect potential intermediates in the catalytic cycle.

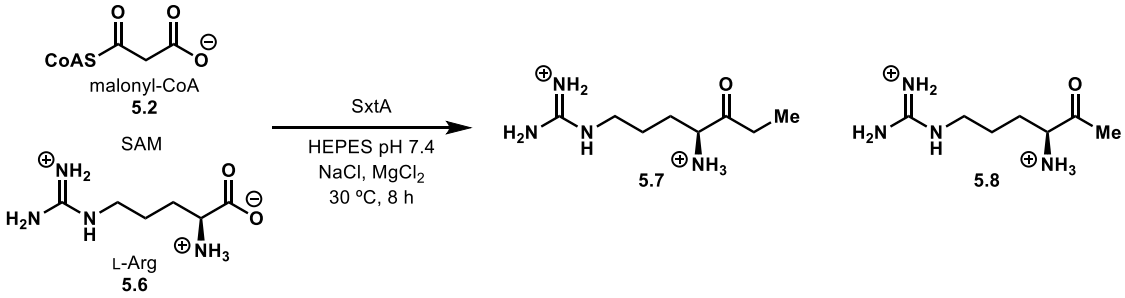
5.2 SxtA MT Structure and Mutagenesis

We first observed unusual SxtA MT activity during experiments to characterize the *C. raciborskii* SxtA module (see Chapter 2.3). The preliminary reactions containing SxtA module, malonyl-CoA, SAM and arginine omitted the $MnCl_2$ salt. When the reaction mixture was analyzed by LC-MS, we observed a 13:87 distribution of the two ketone products **5.7** and **5.8** (Table 5.1, entry 1). Since 87% of the ketone formed had bypassed the MT domain, we speculated at the time that it was simply a very slow-acting domain. Under these conditions, malonyl-CoA (**5.2**) was the only CoA substrate that showed that the MT domain had been active, suggesting that malonyl-CoA was the SxtA starter unit and malonyl-ACP the MT substrate.[3]

Dr. Meredith Skiba also arrived at the same answer from an X-ray crystal structure of AprA MT_L, an MT adjacent to an inactive GNAT domain from the apratoxin A biosynthetic pathway (48% sequence identity to SxtA MT-GNAT).[4,5] The *apo* structure of AprA MT_L (PDB: 6B39) is structurally related to two MTs which methylate the α-carbon of 1,3-dicarbonyl substrates (see

5.9, Figure 5.2),[6,7] and a third enzyme, MppJ, an Fe^{3+}-dependent *C*-MT that methylates its substrate (**5.11**) adjacent to a 1,2-dicarbonyl moiety.[8] Both carbonyl groups were engaged in interactions with the metal center in MppJ; the analogy of this in AprA MT_L and SxtA MT is malonyl-CoA/-ACP coordinated to a metal ion.

Table 5.1. Metal salt screen of SxtA-catalyzed ketone formation



Entry	Metal Salts	% Native Product (5.7)	% Shunt Product (5.8)
1	$MgCl_2$ only	13	87
2	No $MgCl_2$	8	92
3	$MgCl_2$ and $Fe(NH_4)_2(SO_4)_2$	12	88
4	$MgCl_2$ and $MnCl_2$	41	59
5	$MgCl_2$ and $CoCl_2$	4	96

When $Fe(NH_4)_2(SO_4)_2$ salts were added to reactions with AprA MT_L, mono- and dimethylation of malonyl-ACP was observed.[4] However, the addition of iron salts to reactions with the SxtA module had no effect on the ratio of full length (**5.7**) to truncated (**5.8**) ketone products (Table 5.1, entry 3). In a screen of other metal salts, only the addition of $MnCl_2$ increased the MT activity to a 41:59 ratio of **5.7** to **5.8** (entry 4). No dimethylated product (i.e., an isopropyl ketone) was observed. This indicates that the SxtA module is isolated from *E. coli* with some metal already bound in the MT domain, but without full incorporation. The identity of the specific metal and the occupancy level has not yet been confirmed by inductively coupled plasma mass spectrometry (ICP-MS). Acetyl-CoA/-ACP (see **5.14**), which had previously been proposed as the SxtA MT substrate[9] lacks the second carbonyl group for bidentate binding to metal center and the truncated

ketone **5.8** resulting from an acetyl substrate appears to be an *in vitro* shunt product only. There was no indication that the metal centers are redox-active; rather, they are likely acting as Lewis acids to lower the pK_a at the site of methylation in the substrates.[8]

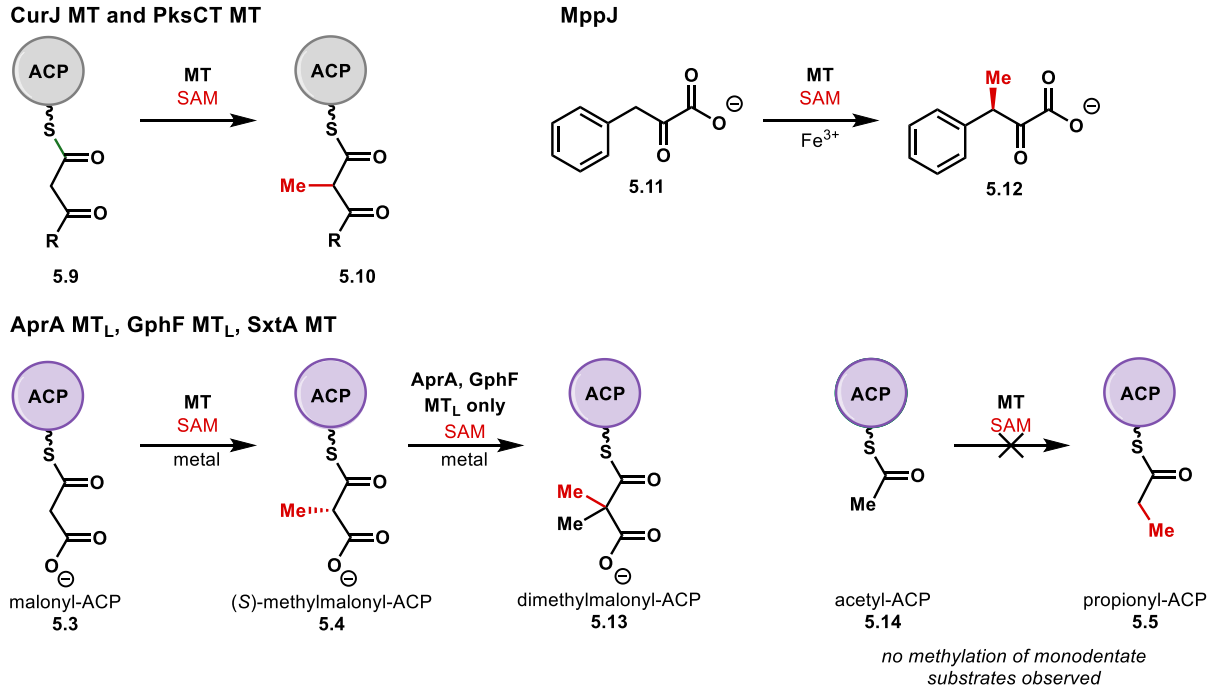


Figure 5.2. The activity of *C*-methyltransferases most closely structurally related to SxtA MT.
CurJ MT,[6] PksCT[7] MT, MppJ,[8] AprA MT_L, GphF MT_L[4] and SxtA MT[3] methylate adjacent to two carbonyl groups.

To identify the metal-binding residues, we aligned the primary sequences and generated a homology model of SxtA MT-GNAT relative to AprA MT_L (PDB: 6B3B).[4] Whereas the Fe^{3+} in MppJ was coordinated by two His residues and two water molecules (see **5.15**, Figure 5.3)[8] and AprA MT_L by His369, His456 and Gln461 (**5.16**), the metal ligands in SxtA MT appear to be His372, Gln459 and Gln464 (Figures 5.3 and 5.4). In addition, the Fe-dependent AprA MT_L and a related enzyme, GphF MT_L, dimethylate malonyl-ACP in their native pathways,[5,10] while SxtA MT with manganese added only monomethylates. However, there was no obvious explanation in

the SxtA MT homology model that would account for the difference in reactivity and or differences in the mechanism of methylation.

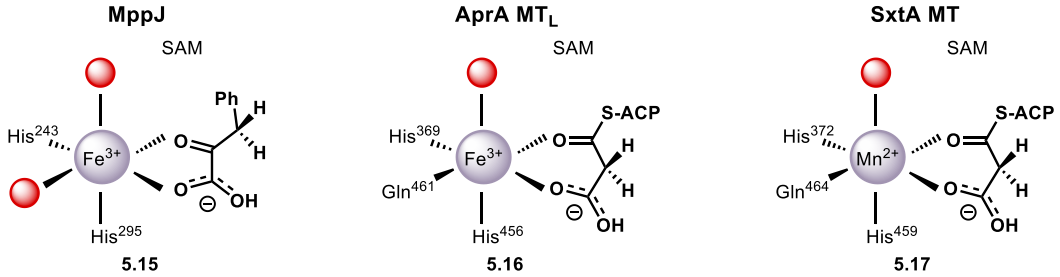


Figure 5.3. Ligands of metal-dependent *C*-methyltransferases.
Red spheres represent water ligands. MppJ PDB: 4KIB;[8] AprA PDB: 6B3B.[4]

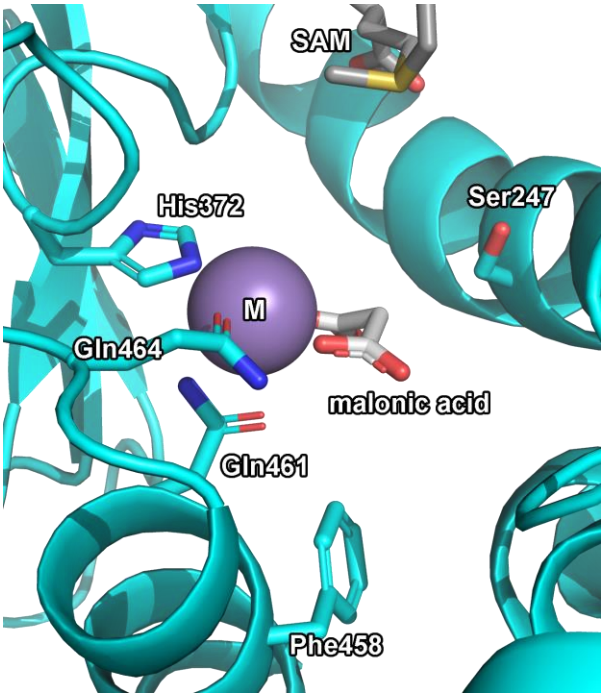


Figure 5.4. Cartoon structure of the SxtA MT homology model.
The SxtA model was based upon AprA (PDB: 6B3B). Potential key residues are labeled. Some residues were omitted for clarity. M: metal ion (Mn^{2+} in the AprA structure and we propose this is also the metal used by SxtA MT).

We had previously proposed that SxtA MT employs a two-electron mechanism because the methylation occurs at an acidic carbon adjacent to two carbonyls. Because the homology model lacks common features of radical enzymes, we anticipate still that SxtA MT catalyzes a two-

electron methylation. With an estimated pKa of approximately 15[11] that is further decreased by coordination to Mn^{2+}, malonyl-ACP could be enolized by a base either above or below the substrate in either the keto or enol forms (see **5.18**, Figure 5.5). The base may be another residue or a coordinated water molecule. In either case, nucleophilic attack by the enolate intermediate on SAM would result in (*S*)-methylmalonyl-ACP (**5.20**). In this configuration, the methyl group is above the metal ion (Figure 5.4), so a second enolization to generate dimethylmalonyl-ACP in AprA/GphF MT_L must be initiated from the bottom face. The second planar enolate intermediate can attack a second SAM molecule above the metal center.

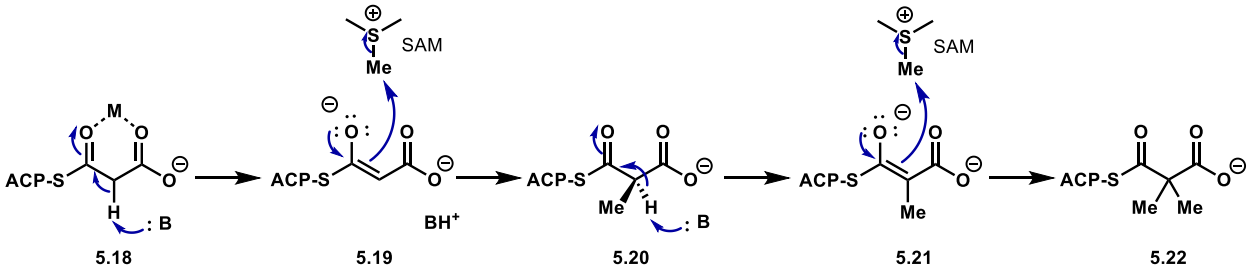


Figure 5.5. Proposed dimethylation mechanism of malonyl-ACP.

In order to investigate if dimethylation can be achieved in SxtA MT as in AprA/GphF MT_L, a series of SxtA MT-GNAT variants was generated. To slow GNAT-catalyzed decarboxylation of malonyl- to acetyl-ACP and allow more time for the MT domain to operate, we first substituted T637 for valine. In addition, Gln459, a metal-coordinating residue, was replaced with His to more closely resemble the AprA MT_L active site. One potential base for malonyl-ACP deprotonation in the SxtA MT active site is Ser247, which suppressed most second methylation activity when substituted with Ala in AprA MT_L (S245A).[4] A third target was Phe458, which is located below the malonyl substrate. This is equivalent to Tyr455 in AprA MT_L, which could function as the base for the second deprotonation event. AprA Y455F also had limited dimethylation activity like S245A.[4]

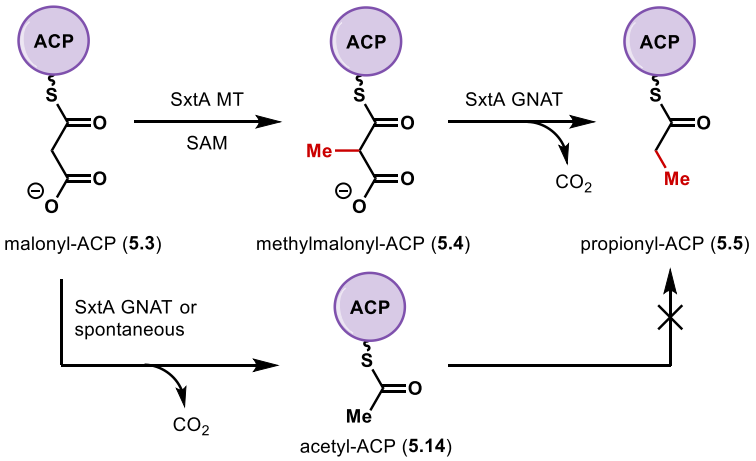


Figure 5.6. Intact protein activity assay of SxtA MT-GNAT activity.

SxtA MT-GNAT variants were incubated with malonyl-ACP for 3 h, and then the mixtures were analyzed by intact protein MS to determine the distribution of acetyl-, malonyl, methylmalonyl- and propionyl-ACP (Figure 5.6). Changing any of the metal-chelating residues significantly affected the methylation activity: SxtA MT-GNAT Q464A and SxtA H372A (see Chapter 2) did not produce methylmalonyl- or propionyl-ACP, and attenuated methylation activity was observed in experiments with SxtA MT-GNAT Q459A/T637V (Figure 5.7). MT-GNAT Q459H/T637V, substitutions intended to match AprA MT_L, attained only 64% methylation compared to 78% methylation in the T637V control reaction. Exchanging the manganese salt for other common metals in additional reactions with SxtA MT-GNAT Q459H/T637V with did not lead to demethylated products (Figure 5.8). We observed protein precipitation in reactions with some divalent metals (Ni, Co, Cu, Zn). Reactions starting from a racemic mixture of methylmalonyl-ACP (**5.4**) were also not methylated a second time. Taken together, these data suggest that only changing the metal and coordinating ligands in SxtA MT is not sufficient to duplicate the dimethylation activity of AprA/GphF MT_L.

Reactions with the variants containing substitutions to the residue which could be acting as a base in the deprotonation of malonyl-ACP, SxtA MT S247A/T637V and F458Y/T637V were very

similar to the T637V control (Figure 5.7, bottom). Methylation was not abolished in the S247A/T637V variant and no dimethylation was detected with the F458Y/T637V variant. These results indicate that Ser247A and Phe458 may not be directly involved in the current monomethylation in SxtA MT. A triple variant, SxtA MT-GNAT F458Y/Q459H/T637V to most closely mimic the AprA MT_L active site was expressed but was removed with the insoluble fraction during purification, so this variant could not be screened. In total, the deprotonation of malonyl-ACP appears to be mediated by residues that we have not found. The possibility that the base may be a water molecule coordinated by unidentified residues has also not been excluded.

Throughout time course reactions with AprA MT_L, the Smith lab observed nearly identical fractions of methylmalonyl- and dimethylmalonyl-ACP (**5.4** and **5.13**) and no buildup of a pool of methylmalonyl-ACP.[4] This suggested that both methylation reactions occur after MT_L binds malonyl-ACP rather than dissociation of methylmalonyl-ACP and a later second methylation event, and was supported by SAM and acyl-ACP binding at opposite ends of the AprA MT_L active site. It is not yet known in SxtA MT if the SAH byproduct can similarly be exchanged for a second molecule of SAM while methylmalonyl-ACP is still associated. The active site volume of SxtA MT is also not known, so it may be possible that interactions between the SxtA MT active site and methylmalonyl-ACP disfavor binding of the bulkier dimethylmalonyl-ACP.

The results from experiments with SxtA MT-GNAT variants indicate that the structural basis of the different levels of methylation between SxtA MT and AprA/GphF MT_L is the sum of many differences from the metal ion identity, metal-coordinating residues, bases and other unidentified factors that may be outside of the first coordination sphere.

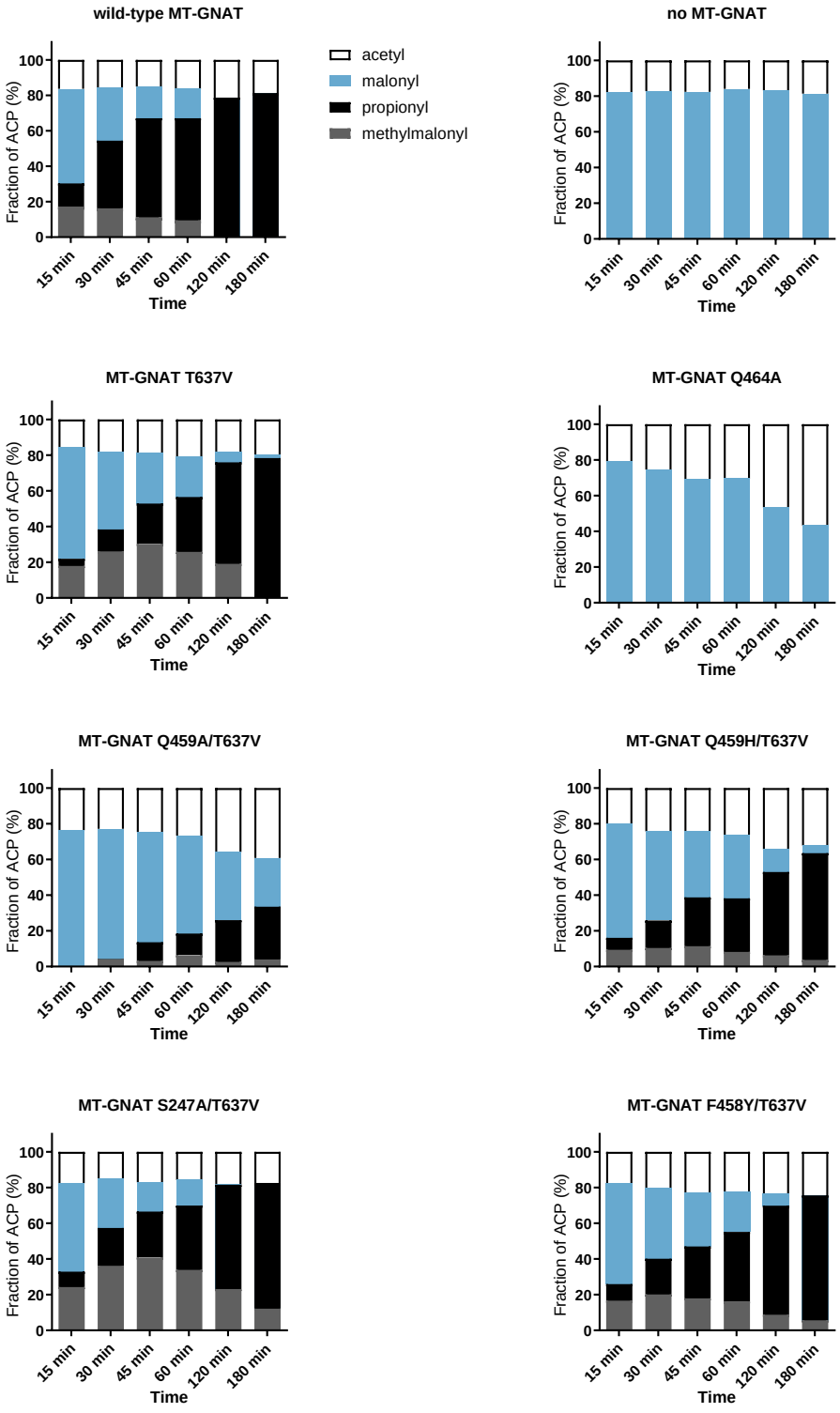


Figure 5.7. Methylation and decarboxylation activity of MT-GNAT variants with mutations in metal-coordinating residues and potential bases to deprotonate malonyl- or methylmalonyl-ACP.

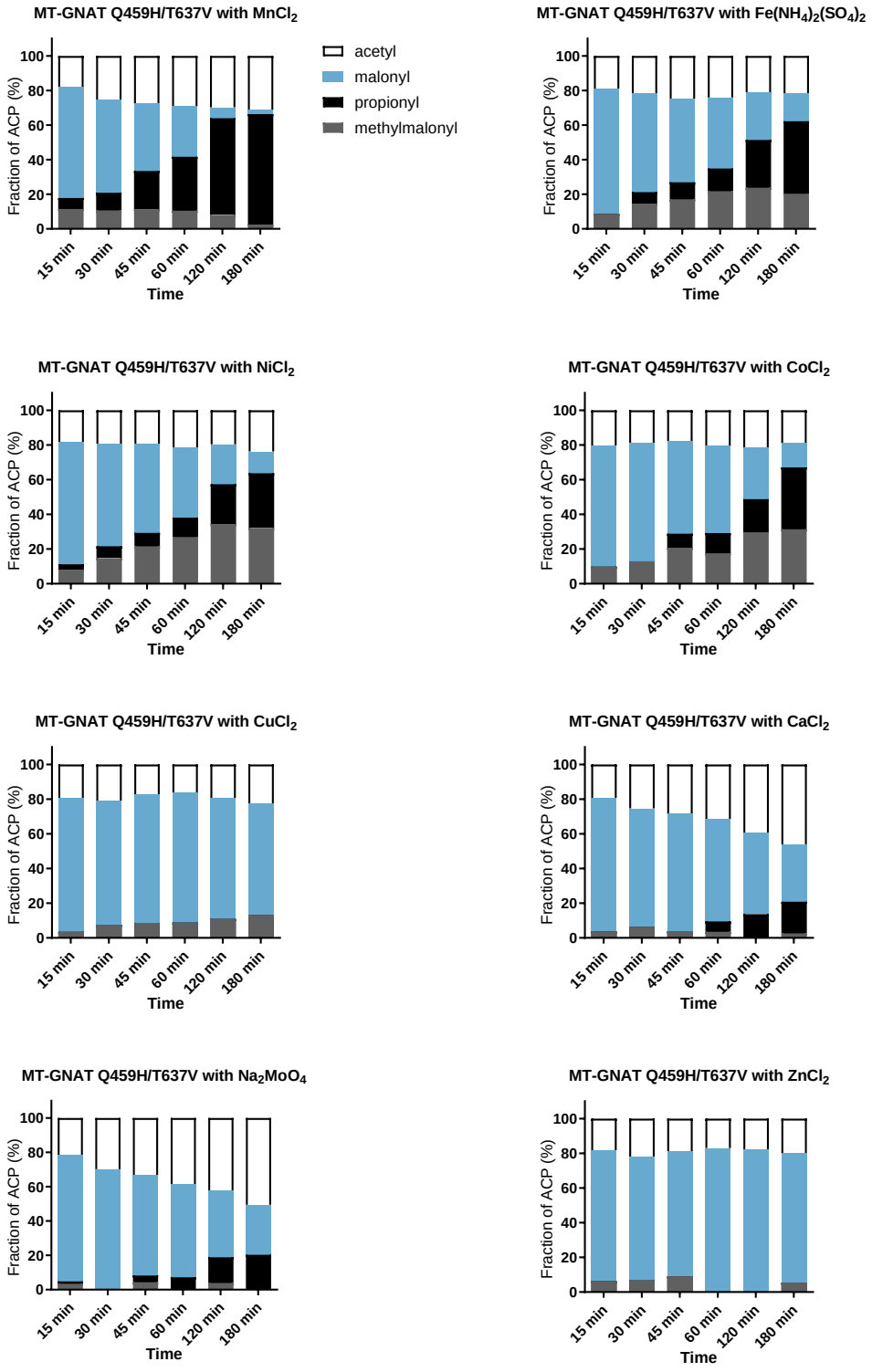


Figure 5.8. Metal salt screen with SxtA MT-GNAT Q459H/T637V.

5.3 UV-Vis Spectroscopy of SxtA AONS

Due to the PLP chromophore, SxtA AONS and some of the PLP-amino acid or PLP-ketone intermediates can be detected by UV-Vis spectroscopy.[2] For these studies, we elected to use the *M. wollei* homolog of SxtA AONS based on its high expression levels and cleanest protein purification.

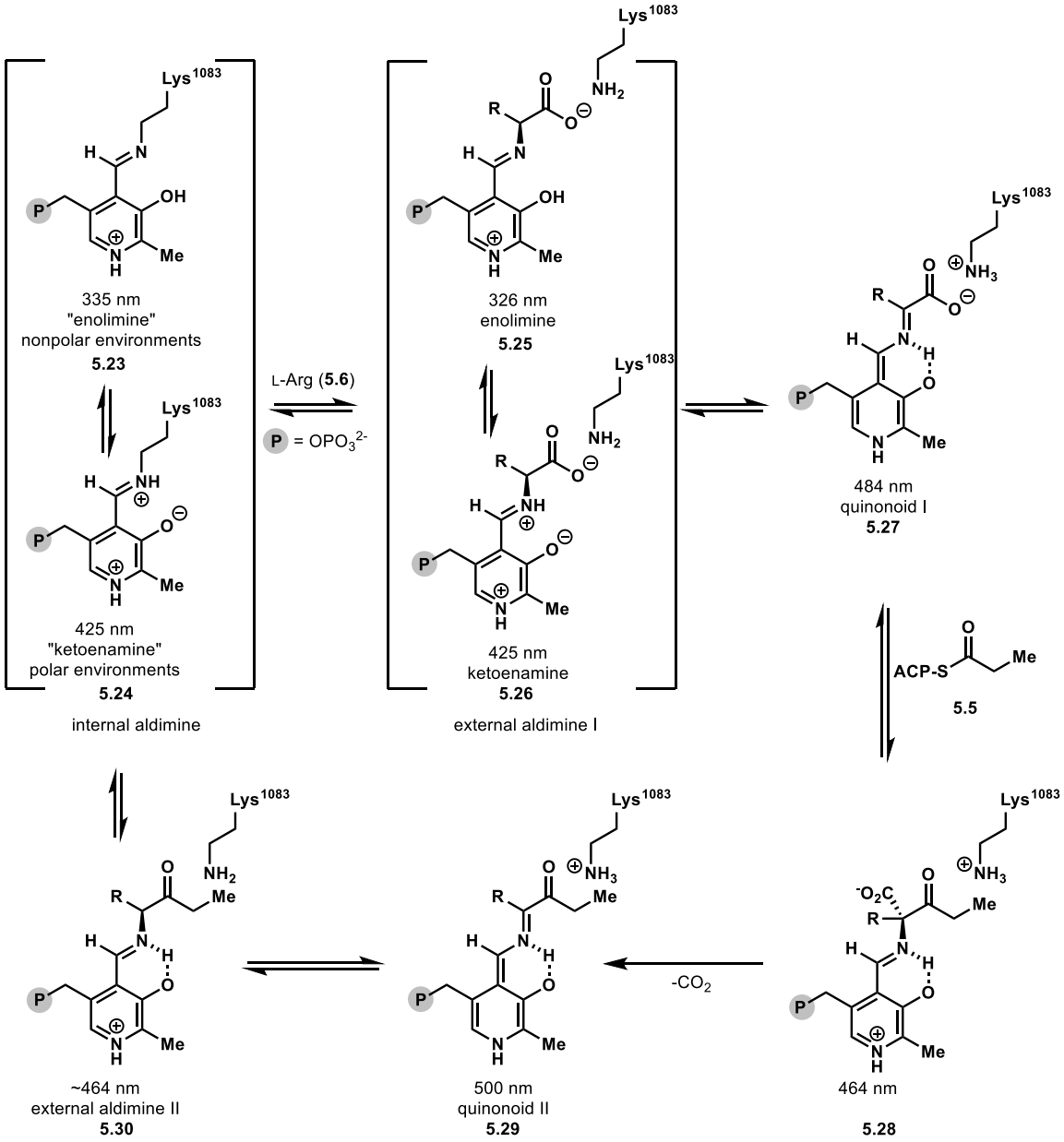


Figure 5.9. Proposed catalytic cycle of SxtA AONS with experimentally determined maximum absorbances of intermediate species.

In its internal aldimine resting state, the PLP-Lys complex in AOS enzymes is in equilibrium between two tautomers: the enolimine intermediate (**5.23**, Figure 5.9) favored by nonpolar environments that absorbs at approximately 330 nm, and the ketoenamine tautomer (**5.24**) that is reported to dominate in polar environments and which absorbs in the 430 nm region.[12] As the amino acid substrate is titrated in, the population ketoenamine tautomer of the external aldimine I intermediate increases in AOS proteins, as we observed with serine palmitoyltransferase (SPT) and L-Ser (Figure 5.10).[13–17] This tautomer further lowers the pK_a of the alpha proton and is proposed to be the more active state.[18]

To observe the tautomeric distribution changes of SxtA AONS with after L-Arg biding, we obtained spectra of the enzyme without substrate and monitored the changes as the amino acid was added. Without substrate present, SxtA AONS is distributed between the enolimine and ketoenamine species (335 and 425 nm, respectively, Figure 5.10), but addition of L-Arg caused the absorbance at 326 nm to increase and flatten at 425 nm, and the enzymatic solution turned from yellow to clear. Although we have determined that quinonoid intermediate I (**5.27**) does form, as indicated by productive deuteration of various substrates (see Chapter 4.2), a new peak at approximately 500 nm corresponding to a quinonoid species has not been observed. Quinonoid **5.27** may be too transient to be seen by UV-Vis, even at higher pH. Spectra were unchanged by the later addition of propionyl-pantetheine (**5.32**), a thioester accepted by SxtA AONS to form the native ketone product **5.7**.

There was little change to the spectra when other substrates, such as D-Arg or L-Arg-OMe, were titrated into the enzyme solutions at pH 7. At basic pH in the presence of L-Arg-OMe, a small peak at 484 nm was observed, which we attribute to quinoinoid I. Additionally, in the presence of propionyl thioester **5.32**, a new peak was observed at 454 nm, likely corresponding to the tetra-

substituted intermediate **5.33**, which cannot decarboxylate until the ester group spontaneously hydrolyzes to a carboxylic acid (Figure 5.11).[19]

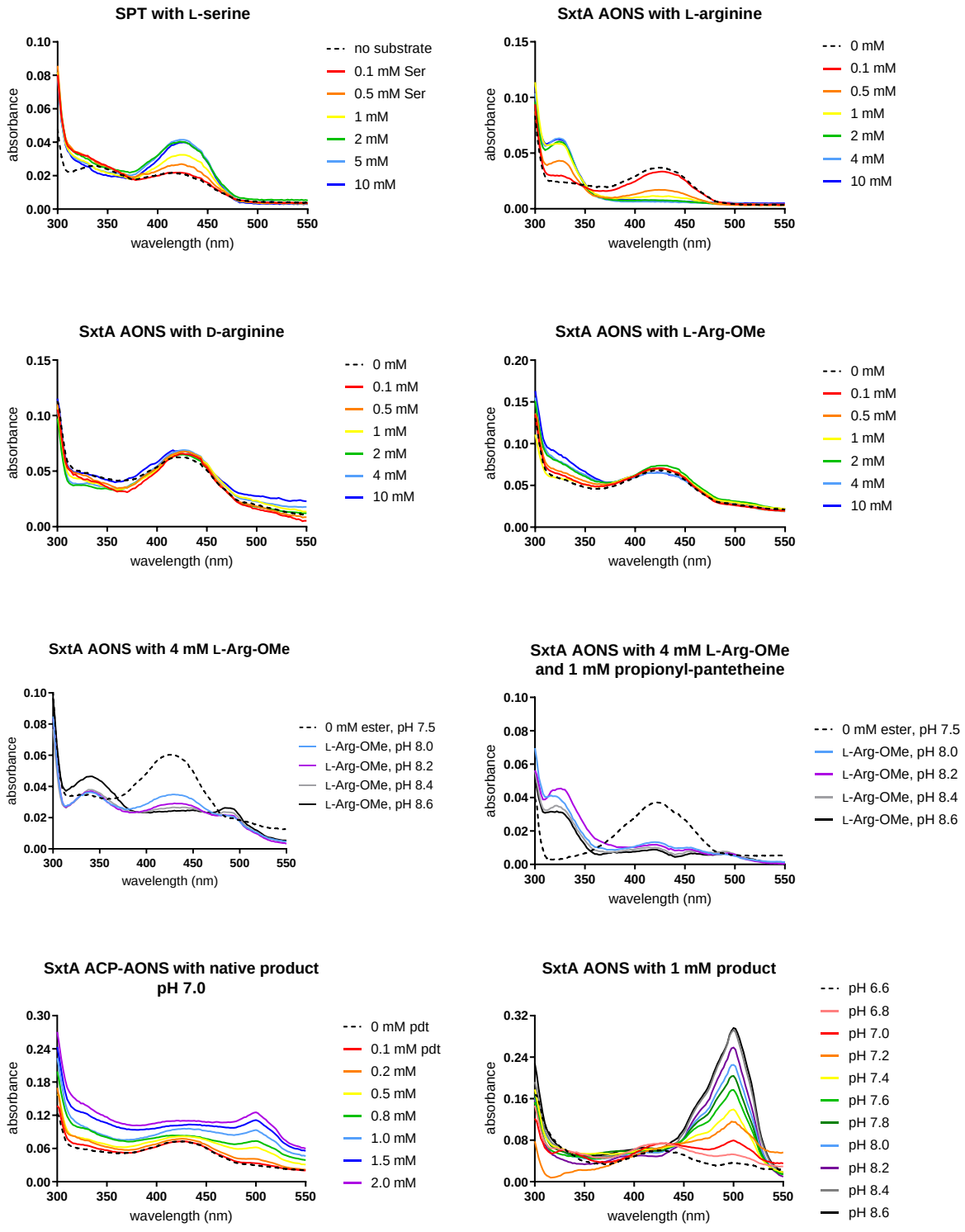


Figure 5.10. UV-Vis spectroscopy of SPT and SxtA AONS under multiple conditions. SPT: serine palmitoyltransferase.

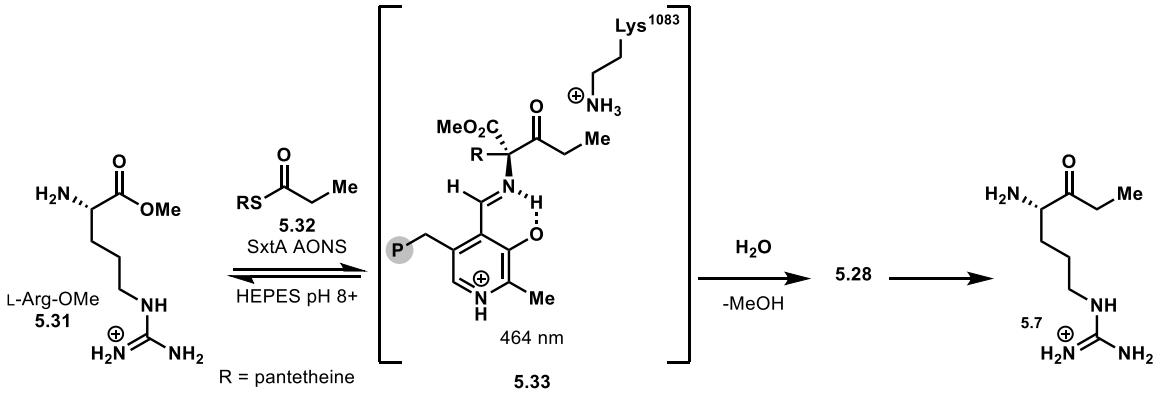


Figure 5.11. Formation of the tetra-substituted PLP-complex from L-Arg-OMe on pathway to the native product.

When the synthetic standard of ketone **5.7** was added, a second quinonoid (**5.29**) appeared at 500 nm, in addition to a small shoulder around 464 nm which may be the second external intermediate (**5.30**). The enzymatic solution turned from yellow to orange to reflect the new maximum absorbance. As the pH increased, the 500 nm quinonoid II peak also grew dramatically. Through spectroscopy of more exact titration and pH shifts, the K_d of the product ketone and pK_a of the ketone-PLP complex can be experimentally calculated. From the procedures of Ferreira and coworkers, we roughly estimate a low micromolar K_d and a pK_a between 8-9 for ketone **5.7**.[20]

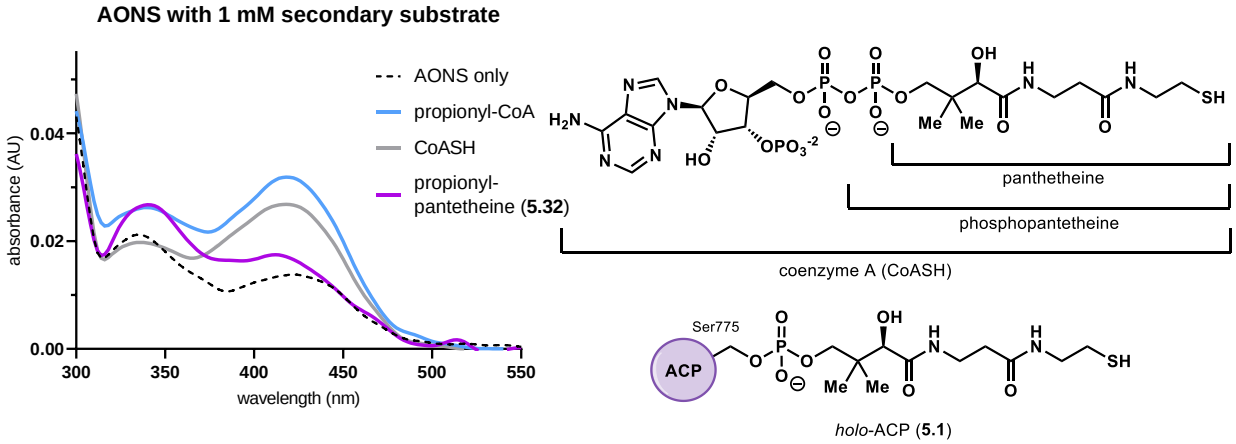


Figure 5.12. Shifting the equilibrium of PLP tautomers in the presence of thiols and thioesters.
Spectra courtesy of Sarah Ackenhusen.

In data obtained by Sarah Ackenhusen, addition of propionyl-CoA and CoASH without L-Arg present shifted the equilibrium of the resting state tautomers toward the active ketoenamine species. The K_d of L-Arg was also determined to be 0.9 mM, similar to the values for other AOS enzymes.[20–22]

These spectroscopic experiments contribute more to our understanding of SxtA AONS. The resting state of the enzyme shows the expected equilibrium between two tautomers (335 and 425 nm, respectively). In contrast to well-studied AOS enzymes, when the native amino acid substrate, L-Arg, binds, the less active enolimine species (326 nm) became the major peak in SxtA AONS. This pattern is the opposite of what has been observed in other reported AOS studies and suggests that the active site is now more neutral after binding of the positively charged Arg.[18] Binding of the bulkier thiols and thioesters (e.g., CoASH and propionyl-CoA) shifts the equilibrium back toward the more acidic ketoenamine tautomer (also 425 nm, see Figure 5.12) to favor deprotonation and formation of the quinonoid intermediate (484 nm, **5.27**). This species is too short-lived to be observed with an L-Arg substrate but can still form in the absence of a thiol or thioester in solution. After addition of the propionyl group and decarboxylation (**5.28** to **5.30**, Figure 5.9), quinonoid II (500 nm) may be reprotonated, and the final ketone product displaced. Because ketone **5.7** likely binds much more tightly to SxtA AONS than L-Arg (Figures 5.10 and 5.12), the rate-limiting step in the AONS catalytic cycle may be product release, as suggested by Ferreira and coworkers.[20,23,24] Basic reaction conditions favor quinonoid II over external aldimine II, which may elucidate why we previously observed a decline in overall SxtA AONS activity as the pH increased (Chapter 3.2).

We attempted to test for product inhibition by incubating SxtA AONS with unlabeled synthetic ketone **5.7** then adding ^{15}N-arginine and propionyl thioester **5.32** to synthesize labeled ketone

(Figure 5.13). The quantification was highly variable due to the storage of quenched samples, and the total ketone present did not converge on one value. The possible inhibition to explain the relatively short AONS condensation lifetime was not very dramatic. It may be a combination of factors, not limited to product inhibition that leads to a cessation of AONS condensation activity in under 4 h.

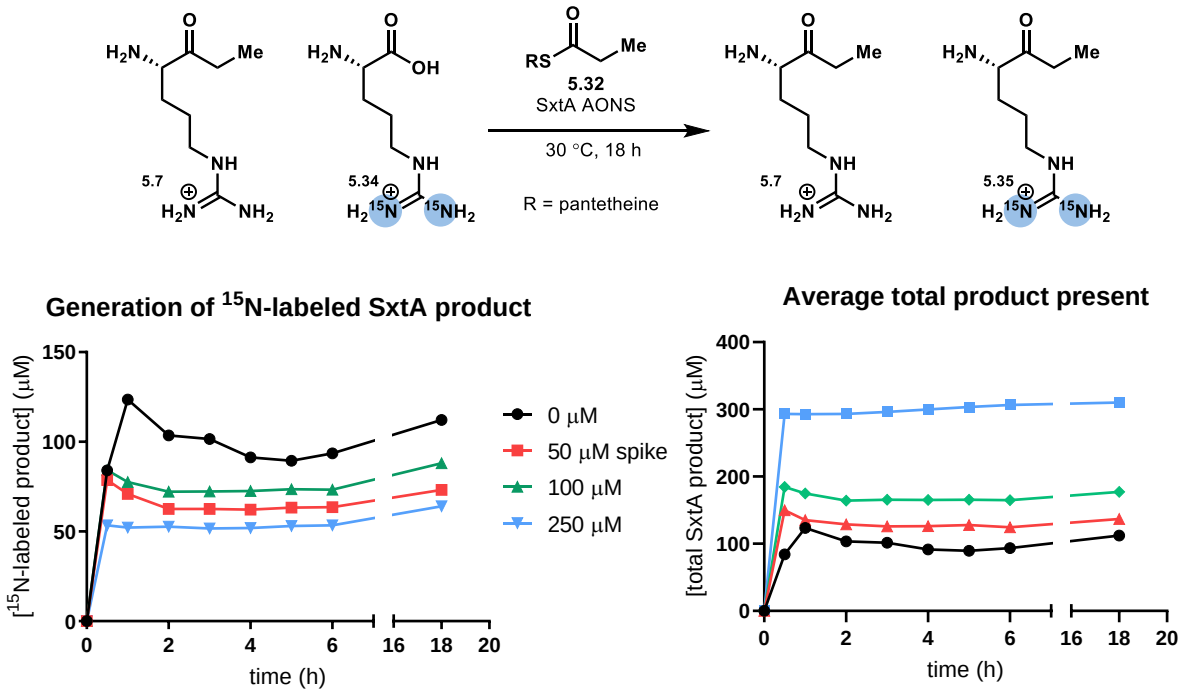


Figure 5.13. Biosynthesis of ^{15}N-labeled ketone with unlabeled synthetic product spiked in. The addition of synthetic standard **5.7** ranged from 0-250 μM.

To improve SxtA AONS as a biocatalyst for formation of amino acid derivatives, we may engineer the active site to favor the currently minor but more active ketoenamine tautomer by introducing more polar side chains. Additionally, some interactions between the enzyme and the acyl group of the thioester partner substrate/product ketones could be disrupted to weaken the binding of ketones and promote product dissociation.

5.4 Conclusions

We elucidated SxtA's native function previously within saxitoxin biosynthesis and then proceeded to probe possible non-native activities of two partner enzyme domains, MT and AONS. Specifically, for the MT domain, we investigated the possibility of introducing dimethylation activity, and for SxtA AONS, improvement as a general biocatalyst for the modification of amino acids.

We had deduced the starter unit of SxtA and the substrate of the MT domain from activity tests of a panel of acyl-CoAs and matching the final module products, ketones **5.7** and **5.8** to synthetic standards. Our suspicions were confirmed by structural data of AprA MT_L from Dr. Meredith Skiba and coworkers showing that GNAT-associated MTs like SxtA MT are metal dependent. Mn^{2+} was the only metal that increased SxtA MT activity, improving the ratio of native to shunt product from 13:87 to 45:55. Attempts to determine the exact residues responsible for the deprotonation and thus the difference between the monomethylation activity of SxtA and dimethylation by AprA/GphF MT_L have not been fruitful. We may make a more informed evaluation of MT mechanism is when a complete crystal structure is obtained.

Although we have observed that SxtA AONS catalyzes many non-native reactions in α-amino ketone generation, a short enzyme lifetime (<4 h) has hampered efforts to run preparative-scale reactions. Employing the corresponding methyl esters of the amino acids (Chapter 4.3) may have caused a conformational change that allowed all esters to be deprotonated and form quinonoid intermediates (Figure 5.14). However, in condensation reactions consisting of SxtA AONS, propionyl-panthetheine and Me esters, we did not observe the formation of new ketone products. It is possible that these reactions are stalled at the tetra-substituted intermediate **5.33**. Due to the known hydrolysis of Me esters to carboxylic acids within AOS active sites (i.e. **5.33** to **5.28**),[19] it is more likely there was no acyl addition to quinonoid **5.36** in Figure 5.14. We previously

suggested that quinonoid **5.36** may be more similar to quinonoid II (**5.29**), to which we have never observed the second addition of an acyl group to generate a diketone product. However, the spectroscopic studies in Chapter 5.3 show that quinonoids **5.29** (maximum absorbance at 500 nm) and **5.36** (484 nm) are spectroscopically distinct. From the binding studies in this Chapter demonstrating that propionyl-CoA but not -pantetheine (**5.32**) shifts the equilibrium toward the more active enzyme state, it is also possible that **5.32** was not enough to also cause a conformational change where the acyl group can add to quinonoids **5.27** or **5.36** at neutral pH.

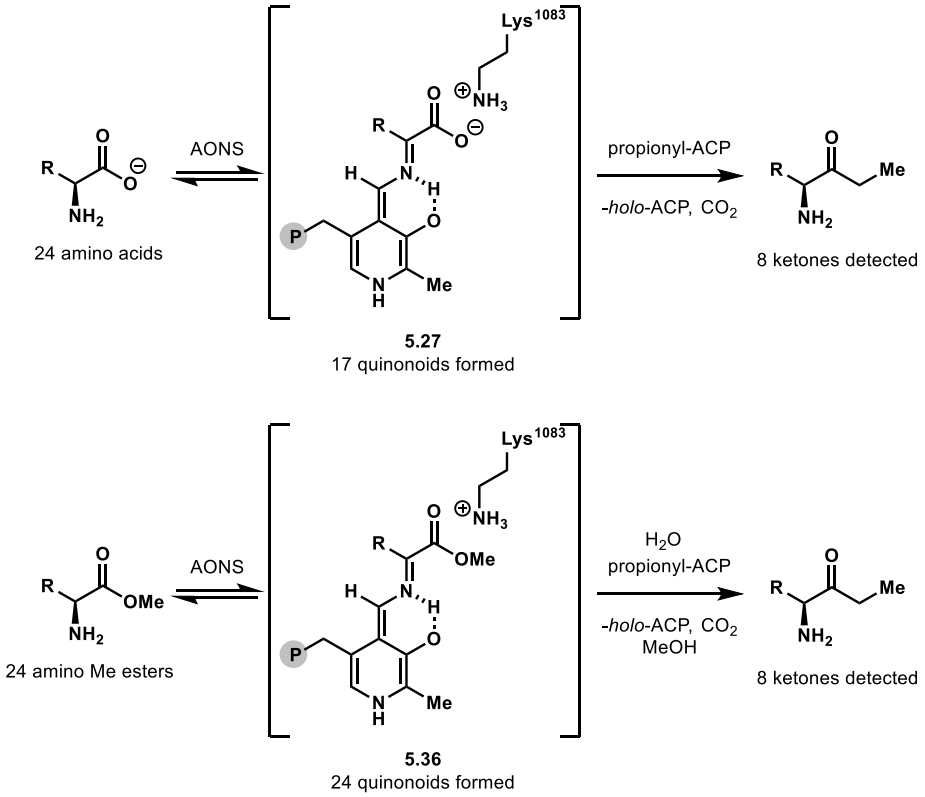


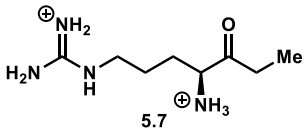
Figure 5.14. SxtA AONS-catalyzed transformation of a panel of α-amino acids and their methyl esters to α-amino ketones.

Based on our results with SxtA and its domains, future work will strive to identify the residues that govern the selectivity of binding different amino acid side chains and interactions between Ppant/CoA/*holo*-ACP and SxtA AONS. These studies may uncover the enzyme conformational changes to allow formation of PLP-tetra-substituted species such as **5.33** and their released

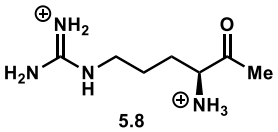
products. On the pathway to α-amino ketones, favorable decarboxylation of **5.28** leads to the second quinonoid **5.29**, and we will also work to increase the enzyme turnover by weakening the enzyme-ketone binding for product release. Potential product inhibition can also be addressed by designing enzymatic cascades to remove ketone **5.7** and raise overall flux. We anticipate that these studies will contribute to the improvement of SxtA AONS as a biocatalyst for preparative-scale synthesis of α-amino ketones, more broadly, for the modification of α-amino acids.

5.5 Experimental

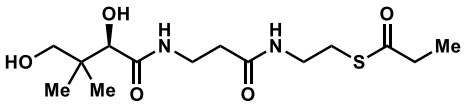
I. Chemical synthesis



5.7



5.8



5.32

Synthetic ketone products **5.7**, **5.8** were synthesized according to the procedures described in Chapter 2. Propionyl-pantetheine (**5.32**) and other arginine-based ketones were synthesized according to the procedures described in Chapter 3. Compounds were stored as 50 mM stock solutions in DMSO.

II. Cloning, protein expression and purification

A. General information

Escherichia coli cloning strain DH5α (Invitrogen) was used for DNA propagation. Protein was expressed in *E. coli* strains BL21(DE3) (New England BioLabs) or BAP1 after transformation with the appropriate plasmids. Cells were grown in Luria Bertani (LB) or Terrific broth (TB) with 4% v/v glycerol supplemented with kanamycin (50 μg/mL) or ampicillin (100 μg/mL) purchased from Gold Biotechnology.

Primers were purchased from Integrated DNA Technologies. DpnI restriction enzyme was purchased from New England BioLabs. Ligation-independent cloning (LIC)-qualified T4 DNA polymerase was purchased from EMD Millipore. QIAquick PCR purification, gel extraction and miniprep kits were purchased from Qiagen. HisPur nickel-nitrilotriacetic acid (Ni-NTA) resin was purchased from Thermo Scientific. Proteins were concentrated using Amicon centrifugal filters purchased from EMD Millipore at 4,000 x *g*, 4 °C. PD-10 desalting columns were purchased from GE Healthcare. Protein samples were analyzed on Mini-PROTEAN TGX Gels (4-15%) from

BioRad and visualized with Protein Ark Quick Coomassie Stain from Anatrace. Proteins were quantified with the Pierce 660 nm Assay Reagent from Thermo Scientific.

pET28b-*sxtA* from *C. raciborskii* was obtained from the laboratory of Prof. Brett Neilan (University of Newcastle) and expressed and purified as described in Chapter 2. pMCSG7-*sxtA*_ACP, pMCSG7-*sxtA*_MT-GNAT, pMCSG7-*sxtA*(T637V) from *C. raciborskii* were subcloned as described in Chapter 2. pMCSG7K-*sxtA*_AONS from *M. wollei* was subcloned, expressed and purified with pET28b-*SPT* as described in Chapter 4.

pET24b-*sfp* was kindly donated by the lab of Professor David Sherman at the University of Michigan, and expressed and purified as described in Chapter 2.

B. Sequence information

pMCSG7 *C. raciborskii sxtA* MT-GNAT (Met1-Ser721), gDNA sequence. GenBank accession ABI75094.1; linkers and the 6x-His tag are shown in bold.

atgcaccatcatcatcatcattcttctggtgtagatctgggtaccgagaacctgtacttccaatccaatgcaATGTT
ACAAAAGATTAATCGTTATACTCACGGCTTTGTGGCGGTTCCCGTTATTCTTGCGTGTCGAGAAAAAGGTGTTTTTG
AATTACTCGCCGATGAAAGTCCTCTCTCTTTAAACCAAATGGTGGAACATCTGGGAGCTAACAGCGGACATTTTCAA
GTTGCTTTGAGGATGCTCGAGTCTTTACATTGGCTTTCCCGAAATAAGGAGCTTAAATACTCTCTAACCGCAGAAGC
AGCGATTCACAACAAAATTTCGGAAGACATTCTTCAATTGTACAACCTACCAATTCAGTCTTATTTAGAAGGGAAAC
AAGGAAATTTGCTGGGAAGATGGATTGAGCGTTCTTGCCAATTGTGGAACCTGGACAATCCCCTAATGGCAGATTTT
TTAGATGGATTACTGGTCATCCCACTCCTGCTGGCACTGCACAAACACAACTTGCTTGCAGATTCGGAGGACAAACC
TTTGCTCTCCTCATTAAGCTCAACAGTGCAAGAAGAGTTGGGTAAGTTATTTCTCCACCTTGGCTGGGCTGACCTTA
CAGCAGGTCGTTTGACCATAACCGAACTTGGTCGATTTATGGGAGAGCGAGCCTTGAATACAGCCATAGTGGCGTCC
TACACTCCTATGTTGTCCCGCATTCATGATGTATTGTTTGGCAATTGTCTCTCCGTATTCCAAAGAGATGCTTCCGG
TCACGAAAGGCACATTGATCGCACCCTTAACGTGATCGGGAGTGGATTTCAACACCAGAAATATTTTGCCGATTTAG
AAGAAAGTATCCTCTCAGTATTCAATCAGTTGCCATTAGAAGAACAACCCAAATACATTACTGACATGGGGTGTGGC
GATGGAACTCTCCTGAAACGAGTCTGGGAAACCATTCAATTTAAGTCTGCTAGGGGAAAAGCACTCGAACAGTATCC
CCTGCGTCTTATAGGTGTAGATTATAACGAAGCTTCTTTAAAAGCTACCACACGCACCCTTGCTAGCCTTCCCCACT
TAGTTTTACAGGGAGATATTGGGAACCCAGAACAAATGGTGCGTTCTTTAGAAGCTCATGGCATTCATGATCCCGAA
AATATCCTGCACATCCGTTCGTTCCTCGACCATGATCGTCTCTTTATTCCTCCTCAGAAAAGAAACGAATTGAAAGA
ACGTGCTCACTTACCTTACCAATCAGTCTGTGTCGATGATCAAGGAGAGCTTATTCCTCCTCATGTTATGGTGCAAA
GTTTGGTGGAACACTTAGAAAGATGGTCTCAAGTGGTCAATAAACACGGTTTAATGATTTTGGAGGTCCATTGTTTG
GAACCAAGGGTAGTCTATCAGTTTTTAGACAAAAGCGAAAACTTACATTTCGATGCGTTTCAGGGATTTTCTCAGCA
GTATCTTGTGGAAGCTGAGGTTTTTCTCATGTCTGCTGCACAAGTAGGTCTATTTCCAAAACTAGAGCTTTCTAAAA
GATACCCAAAAACATTTCCTTTTACTCGCATTACGCTTAATTACTTCGAGAAAAGACCTTACAAAATTAGTCATGCC
TATTTGTCAGATTTACCTGCCTTAGTTGACTTGGAGGTCAAGTGTTGGCCAGAAAATTTACGGGCATCTACTCATGA
AATTCGGCGAAGACTTGAGCTAAACCCGCAAGGTAATTTAGTGCTGATTATAGAAGATCAAATTATTGGTGCGATTT
ATTCCCAAACAATTACCAGCACTGAGGCATTAGAGAATGTAAAATATGCGCAAGTGCCGACGTTACATACTCCCCAA
GGGTCAGTTATTCAACTGCTCGCACTAAATATTCTACCTGAGTTTCAGGCGCGGGGGTTAGGAAATGAATTGCGGGA
CTTTATGCTTTACTACTGTACCCTGAAAGGCGGCATTGAGAGCGTGGTGGGTGTAACTCGCTGTCGAAATTATGTCA
ATTATTCCCAAATGCCGATGATGGAGTATTTAAAGCTACACAATGAGCAACGACAGCTTCTGGATCCAATTGTGGGT
TTCCATGTGTCGGGAGGAGCCGAAATTAGGGGAATTATTGCTAATTATCGTCCGGAAGATACAGATAATCTCGGCAT
GGGTATTTTGATTGAATATAACCTGCGCGACAGTGCTTTGCACTCGCCTGGTGATCGCAAAGGACCGTATATTAACT
CA

Translation:

MHHHHHHSSGVDLGTENLYFQSNAMLQKINRYTHGFVAVPVILACREKGVFELLADESPLSLNQMVEHLGANSGHFQ
VALRMLESLHWLSRNKELKYSLTAEAAIHNKISEDILQLYNLPIQSYLEGKQGNLLGRWIERSCQLWNLDNPLMADF
LDGLLVIPLLLALHKHNLLADSEDKPLLSSLSSTVQEELGKLFLHLGWADLTAGRLTITELGRFMGERALNTAIVAS
YTPMLSRIHDVLFGNCLSVFQRDASGHERHIDRTLNVIGSGFQHQKYFADLEESILSVFNQLPLEEQPKYITDMGCG
DGTLLKRVWETIQFKSARGKALEQYPLRLIGVDYNEASLKATTRTLASLPHLVLQGDIGNPEQMVRSLEAHGIHDPE
NILHIRSFLDHDRLFIPPQKRNELKERAHLPYQSVCVDDQGELIPPHVMVQSLVEHLERWSQVVNKHGLMILEVHCL
EPRVVYQFLDKSENLHFDAFQGFSQQYLVEAEVFLMSAAQVGLFPKLELSKRYPKTFPFTRITLNYFEKRPYKISHA
YLSDLPALVDLEVKCWPENLRASTHEIRRRLELNPQGNLVLIIEDQIIGAIYSQTITSTEALENVKYAQVPTLHTPQ
GSVIQLLALNILPEFQARGLGNELRDFMLYYCTLKGGIESVVGVTRCRNYVNYSQMPMMEYLKLHNEQRQLLDPIVG
FHVSGGAEIRGIIANYRPEDTDNLGMGILIEYNLRDSALHSPGDRKGPYINS

pMCSG7-*C. raciborskii sxtA* ACP genomic DNA sequence (Ala722-Gln815)

atgcaccatcatcatcattcttctggtgtagatctgggtaccgagaacctgtacttccaatccaatgcgatgGC
AATTGGTTCATTGGTACCAAAAGCAACATCTGCAACTAAGGAAAACAAAACTGTAGCGGATCTCGTTAAAGAATGCA
TCTTAAAAGTAATGGGTTCCCAACGTCAGGCAGCCTACGCTCCACAACAAAAACTGCTGGATATGGGATTAGATTCT
TTAGATTTATTAGAACTGCAAACGCTCCTAGAGGAACGTTTAGGGATCAATCTGTCTGGAACGTTCTTTTTACAAAA
GAACACTCCAACTGCCATCATCACTTATTTCCAAAACCAAGTGGTACAA

Translation

MHHHHHHSSGVDLGTENLYFQSNAMAIGSLVPKATSATKENKTVADLVKECILKVMGSQRQAAYAPQQKLLDMGLDS
LDLLELQTLLEERLGINLSGTFFLQKNTPTAIITYFQNQVVQ

pMCSG7K *Microseira wollei sxtA* AONS *E. coli*-optimized sequence

GenBank accession ACG63826.1; linkers and the 6x-His tag are shown in bold.

atgcaccatcatcatcattcttctggtgtagatctgggtaccgagaacctgtacttccaatccaatgcaGAAAA
GCAGAGTGATCTGGCGAGCCCCGTGGACTTAGCTAATGAAATCAACACATTAGAAAATGTTGTCAATCAGCAGAAGA
TCCCTCAGGTTACGCGCGTAGTGACTGAACAGCAAGGTCGTAAGGTCTTGATCGATGGCCACTGGGTGATTGACTTC
GCCTCATGCAACTATTTAGGTTTAGACTTGCACCCCAAAGTGAAGGAGGCTATCCCCCCGGCTTTGGAGAAGTGGGG
GGTCCATCCCTCTTGGACACGTTTGGTGGCCAGCCCGGTAATCTATGAGGAATTAGAGGAGGAGTTAAGCAAACTGT
TAGGAGTTCCCGACGTATTAGTTTTCCCAAGCGTCACATTACTTCAGATCGGGGTCTTGCCACTTCTGACAGGCTCG
AACGGAGTTATTTTAGGCGACATTGCTGCGCATCGTTGCATTTACGAGGCCTGCTGCCTGGCACAGCAGAAGGGCTC
ACAGTTTATCCAATATCGCCACAATGACTTGAATGATTTAGCGGAGAAATTGGCAAAGTATCCACTGGAACAAGTCA
AAATCATCGTCATTGACGGAGTGTACTCTATGTCCGCCGACTTCCCAGACCTTCCGGCTTACGTCCGCCTGGCGAAG
GAATACAATGCTTTTATTTATATGGATGACGCGCACGGATTTGGTATTTTGGGTGAGAACCCGTCCTCCGATATGCC
TTACGGATACAAAGGTAATGGTATGGTGAATTATTTCGATCTTCAATTTGCCGAGGACAACATTATCTACGTCGCGG
GGTTGTCTAAAGCCTACAGTTCGTACGCCGCCTTTCTTACTTGCGGTGATCGTCAGATTAAGACCAATTTCCGTAAT
GCTTGGACTGCCATTTTCAGCGGTCCTAGCCCAGTGGCAAGCTTGGCATCAGCCTTGGCCGGCTTACAGGTTAATCG
TCAGGAAGGAGAACAACTTCGCAAGCAAATCTACCACTTGACGCGCAAATTGGTGACCCAAGCTCGCGCCATTGGCT
TCGAAGTGGACAACTACGGTTACGTTCCCATCGTTTCCGTCTTAGTCGGAGACGCGCAACACATGATTGACGTATGT
CAGCTGCTGTGGGAGTACGGCATTCTGATCACACCCGCTATTTTCCCAATTGTTCCACTTAATAAATCCGCGCTGCG
CTTTTCTATCACAGCTGCCAATACGGAAGAAGAGATTGACCAAGCCATCAAAGCATTGGAGGGGGTATGGGTCCTGC
TTCAGAAAGAGAAG**cattag**

Translation: linkers and the 6x-His tag are shown in bold. There is an extra His residue at the C-terminus, added as a mistake in the primer design.

MHHHHHHSSGVDLGTENLYFQSNAEKKSDLASPVDSANEINTLENVVNQQKIPQVTRVVTEQQDRKVLIDGHWVIDF
ASCNYLGLDLHPKVKEAIPPALEKWGVHPSWTRLVASPAIYEELEEELAKLLGVPDVLLFPSLTLLQMGVLPLLTGN
NGVIFGDISAHRCIYEACCLAQHKGAQFIQYRHNDLNDLAAKLAKYPLEQVKIIAIDGVYSMSADLPDLPAYVRLAK
QYNALIYIDDAHGFGILGENPSSDMPYGYKGNGIVNYFDLRFAEDNIIYVAGLSKAYSSYAAFVTCGNSQIKTQFRN
AWTAIFSGPSPVASLASALAGLQVNRQEGEQLRKQVYYLTHKLVTQARAIGFEVDNYSYVPIVSVLVGDAQHIVDVC
QLLWEYGILITPAIFPIVPLNKTALRFSITAANTEEEIDQAIKALEAVWDLLQKRKALLCK**H***

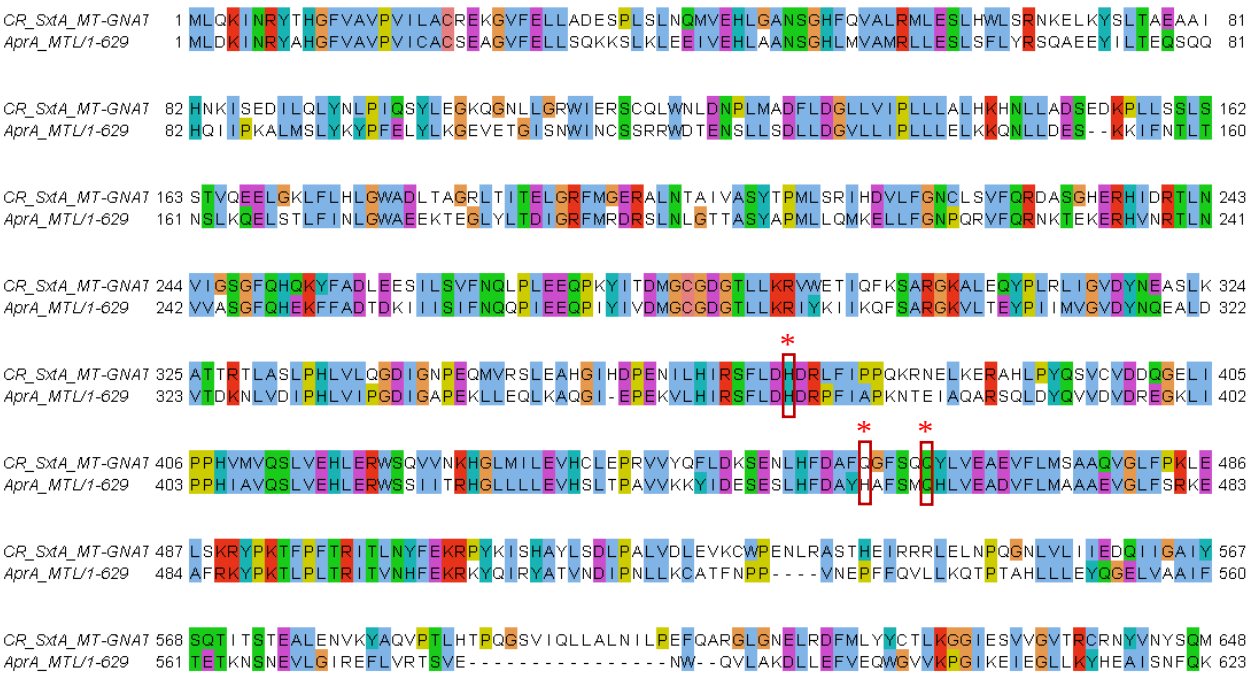


Figure 6.S1. Sequence alignment of SxtA MT-GNAT (1-721) and AprA MT$_L$-ΨGNAT (1-629).

Overall sequence identity is 48%. Sequences were aligned with Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) and displayed in JalView. The metal-binding residues are boxed in red and starred.

C. Cloning

Table 5.S1. Primers used to generate constructs in this Chapter.

Name	Sequence
Primers for site-directed mutagenesis of whole plasmids	
SxtA S247A	5′-CACATTGATCGCACCCTTAACGTGATCGGGGCTGGATTTCAACACCAGAAATATTT-3′
SxtA F458Y	5′-GACAAAAGCGAAAACTTACATTTCGATGCGTATCAGGGATTTTCTCAG
SxtA Q459H	5′-AAAAGCGAAAACTTACATTTCGATGCGTTTCACGGATTTTCTCAGCAG-3′
SxtA Q459A	5′-AAAAGCGAAAACTTACATTTCGATGCGTTTGCGGGATTTTCTCAGCAG-3′
SxtA F458Y/Q459H	5′-GACAAAAGCGAAAACTTACATTTCGATGCGTATCACGGATTTTCTCAGCAG-3′
SxtA Q464A	5′-CATTTCGATGCGTTTCAGGGATTTTCTCAGGCGTATCTTGTGGAAGCTGAGG -3′
Primers for ligation-independent cloning (LIC)	
SxtA MT M1 LIC forward	5′-**TACTTCCAATCCAATGCA**ATGTTACAAAAGATTAATCG-3′
SxtA GNAT S721 LIC reverse	5′-**TTATCCACTTCCAATGCTA**TGAGTTAATATACGGTCCTT-3′

Mutations are highlighted in red. LIC handles in bold.

SxtA MT-GNAT T637V: The excised didomain was amplified from pMCSG7-*sxtA*(T637V) (see Chapter 2) in a 50 µL reaction containing 10 µL Phusion HF buffer, 2 ng/µL parent plasmid, 2 µM each of forward and reverse LIC primers, 200 µM each of dNTPs, 0.04 U/µL Phusion HF. Excised domains were amplified according to the following PCR procedure: 95 °C 2:00, (95 °C 0:30, 50 °C 0:30, 72 °C 0:45/kb) for 30 cycles, 72 °C 10:00. The insert was purified by gel extraction and subcloned into pMCSG7 using standard LIC protocols.[25]

Other SxtA MT-GNAT variants: pMCSG7-*sxtA*_MT-GNAT variants were generated by site-directed mutagenesis of pMCSG7-*sxtA*_MT-GNAT(wt) or pMCSG7-*sxtA*_MT-GNAT(T637V). 50 µL PCR reaction mixtures contained 10 µL Phusion HF buffer, 2 ng/µL wt parent plasmid, 2 µM each of the S773A forward and reverse primers, 200 µM each of dNTPs, 0.04 U/µL Phusion HF and 6% (v/v) DMSO. Amplification was accomplished with the following PCR procedure: 95 °C for 2:00, (95 °C 0:30, 48-60 °C 1:00, 68 °C 4:00) for 18 cycles, with a final extension of 68 °C

for 10:00. This was followed by a 10 µL digestion containing 1 µL NEB CutSmart buffer, 8 µL of the PCR mixture and 20 units of DpnI. The reaction mixture was incubated at 37 °C for 3 h and transformed into chemically competent DH5α cells. Colonies containing the correct mutation(s) were verified by Sanger sequencing.

D. Protein overexpression and purification

Overexpression of all proteins (SxtA MT-GNAT variants, *apo*-ACP, and AONS): pMCSG7 or pMCSG7K plasmids containing the desired insert were transformed into chemically competent BL21(DE3) *E. coli* cells. A single colony was picked to inoculate a 5 mL LB starter culture grown overnight at 37 °C, 200 rpm. The following day, 0.5 L TB media supplemented with either kanamycin or ampicillin, was inoculated with the starter culture and incubated at 37 °C, 250 rpm until the OD_{600} reached 1.0. For *apo*-ACP, the media was also supplemented with a trace metals mixture (final concentrations: 50 µM $FeCl_3$, 20 µM $CaCl_2$, 10 µM $MnCl_2$, 10 µM $ZnCl_2$, 2 µM $CoCl_2$, 2 µM $CuCl_2$, 2 µM $NiCl_2$, 2 µM Na_2MoO_4, 2 µM Na_2SeO_3, 2 µM $B(OH)_3$). Cultures were equilibrated at 20 °C for 1 h. Expression was induced by addition of IPTG (final concentration 200 µM). Cultures were incubated at 20 °C, 200 rpm for 18 h.

Purification of all proteins: Cell pellets were resuspended in 4 mL of lysis buffer (50 mM HEPES, 300 mM NaCl, 10 mM imidazole, and 10% (v/v) glycerol at pH 8.0) per gram of wet cell mass. For SxtA AONS, 1 mM PLP was added to the lysis buffer. Cells were lysed by sonication (3 s on, 6 s off, 5 min total). Insoluble material was removed by centrifugation (30,000 x *g* for 30 min at 4 °C). The clarified lysate was incubated with gentle shaking along with 1-2 mL of nickel-NTA resin for 1 h at 4 °C and poured into a column. The resin-bound protein was washed with 25

mL wash buffer (50 mM HEPES, 300 mM NaCl, 25 mM imidazole, 10% (v/v) glycerol, pH 8.0). The desired protein was eluted with 6 mL elution buffer (50 mM HEPES, 300 mM NaCl, 300 mM imidazole, 10% (v/v) glyercol, pH 8.0). The eluted protein was concentrated using 3 – 50 kDa centrifugal cutoff filters and exchanged into storage buffer (50 mM HEPES, 200 mM NaCl, 10% (v/v) glycerol, pH 7.4) using a PD-10 column. The desalted protein was concentrated further using 10 – 50 kDa centrifugal cutoff filters, aliquoted, flash frozen in liquid nitrogen and stored at -80 °C.

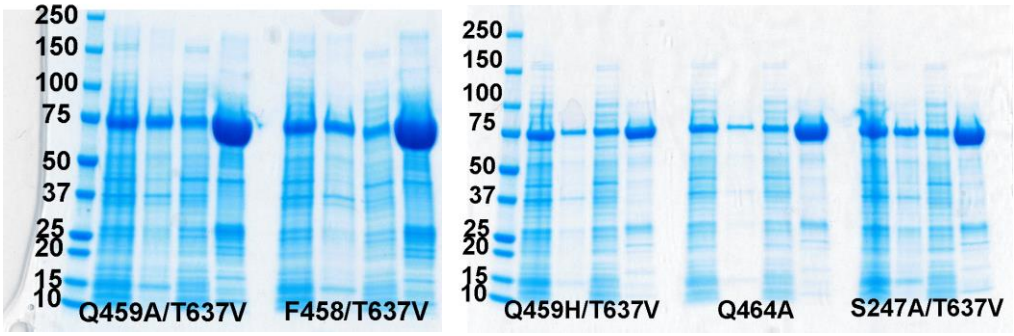


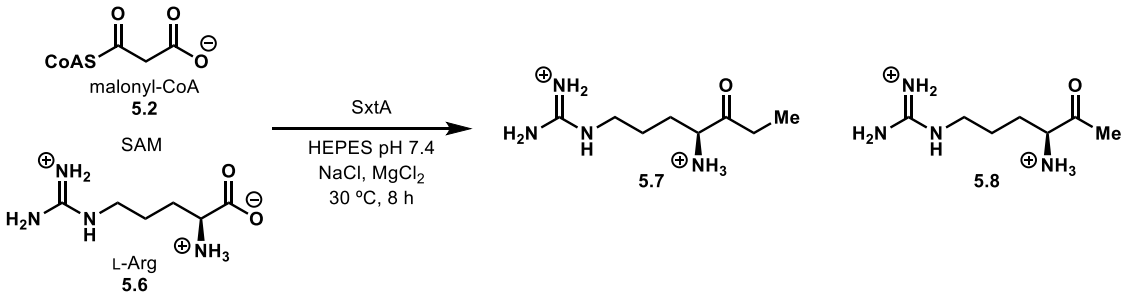
Figure 5.S2. Denatured SDS-PAGE protein gels of new constructs used in this Chapter. SxtA MT-GNAT molecular weight with intact Met1: 84,167.6 Da. Each set of four lanes per construct shows the crude lysate, insoluble material, soluble fraction and isolated protein from left to right.

III. Enzymatic reactions

A. General information

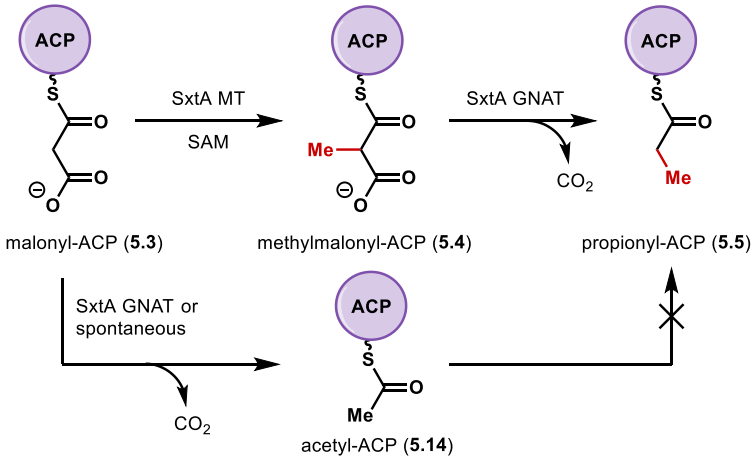
CoASH, propionyl-CoA and malonyl-CoA were purchased from CoALA Biosciences. *S*-adenosyl methionine was purchased as the tosyl salt from Carbosynth and prepared as a 50 mM aqueous solution immediately before use. Electrospray liquid chromatography-mass spectrometry (LC-MS) analysis was performed on an Agilent G6545A quadrupole-time of flight mass spectrometer in positive mode with an Agilent 1290 UPLC system. Solvent A = water with 0.1% formic acid. Solvent B = 95% acetonitrile, 5% water and 0.1% formic acid. MS data were analyzed using the Agilent Qualitative Mass Hunter software. Intact proteins were deconvoluted with the maximum

entropy algorithm. Absorbance spectra of SxtA AONS were obtained on a Shimadzu UV-2501 PC spectrophotometer over a range of 300-550 nm. Malonyl- and methylmalonyl-ACP were prepared as described in Chapter 2.



Metal screen of the SxtA module: 250 μM acyl-CoA, 1 mM SAM, 2 mM Arg, 10 μM SxtA in 50 mM HEPES pH 7.4, 150 mM NaCl, 1 mM $MgCl_2$, 0.5 mM other metal salts (total volume 25 μL) were combined and incubated at 30 °C for 2 h. Reactions were quenched by diluting 5x with acetonitrile. Precipitate was pelleted at 12,000 x *g* for 20 min. The supernatant was diluted 20x with acetonitrile, 1% formic acid, and aqueous 10 μg/mL ^{15}N Arg·HCl. The final ratios (v/v) for analysis were 5% quenched reaction, 2.5% of aqueous 10 μg/mL ^{15}N Arg·HCl (Cambridge Isotope Laboratories), 46% acetonitrile, 46.5% of 1% formic acid in water (e.g., 10 μL of quenched reaction, 5 μL of 10 μg/mL ^{15}N Arg·HCl, 92 μL acetonitrile, 93 μL of 1% formic acid in water for a total volume of 200 μL). Samples were analyzed by LC-MS: column = Waters Acquity 1.7 μm UPLC BEH Amide HILIC 2.1 x 100 mm column; method = 85% B at 0.3 mL/min for 2 min, followed by a linear gradient to 75% B at 0.4 mL/min over 3 min and a second linear gradient to 60% B over 0.5 min, 60% B for 1 min and then 6.5 min re-equilibration at 85% B at 0.3 mL/min (total time 15 min). t_R (**5.7**): 1.94 min; t_R (**5.8**): 2.35 min; t_R (^{15}N Arg): 4.0 min. Ratios were calculated by comparing the areas under the curve of the extracted ion chromatograms. The overall ketone yield was not calculated. Reactions were run in duplicate.

Salts tested: $Fe(NH_4)_2(SO_4)_2$, $MnCl_2$, $NiCl_2$, $CoCl_2$, $CuCl_2$, $CaCl_2$, Na_2MoO_4, $ZnCl_2$. Mn^{2+} was the only salt to show significant improvement over the $MgCl_2$ only control. Fe^{3+}, Ca^{2+} and Mo^{6+} were unchanged from the control, while the remaining salts caused varying levels of precipitation.



Methylation of malonyl-ACP: 100 µM malonyl-ACP and 15 µM MT-GNAT variant in 50 mM HEPES pH 7.4, 150 mM NaCl, 1 mM $MgCl_2$, 0.5 mM $MnCl_2$ (or other metal salt), and 1 mM SAM were combined for a total volume of 35 µL and incubated at 30 °C for 3 h. At 15, 30, 45, 60, 120, and 180 min, 5 µL aliquots were removed and quenched by diluting 12.5x with 1% formic acid in water. Samples were analyzed using a Ppant ejection assay of intact proteins.[26] LC-MS analysis: column = Phenomenex Aeris 3.6 µm WIDEPORE C4 2.1 x 50 mm; method = 5% B at 0.5 mL/min for 2 min, followed by a linear gradient to 100% B over 4 min, 100% B for 2 min, followed by a 0.1 min linear gradient to 5% B and 1.9 min equilibration at 5% B (total time 10 min). t_R = 4.1 min. The relative abundance of ejected Ppant ions for acetyl-, malonyl-, propionyl- and methylmalonyl-ACP were used to calculate the fraction of each acyl-ACP species. They hydrolysis product *holo*-ACP was excluded from the methylation analysis. No dimethylation to dimethylmalonyl-ACP was observed.

Salts tested: $Fe(NH_4)_2(SO_4)_2$, $MnCl_2$, $NiCl_2$, $CoCl_2$, $CuCl_2$, $CaCl_2$, Na_2MoO_4, $ZnCl_2$. Like the full module reactions, Ni^{2+}, Co^{2+} Cu^{2+} and Zn^{2+} salts caused varying levels of precipitation, diminishing MT-GNAT activity.

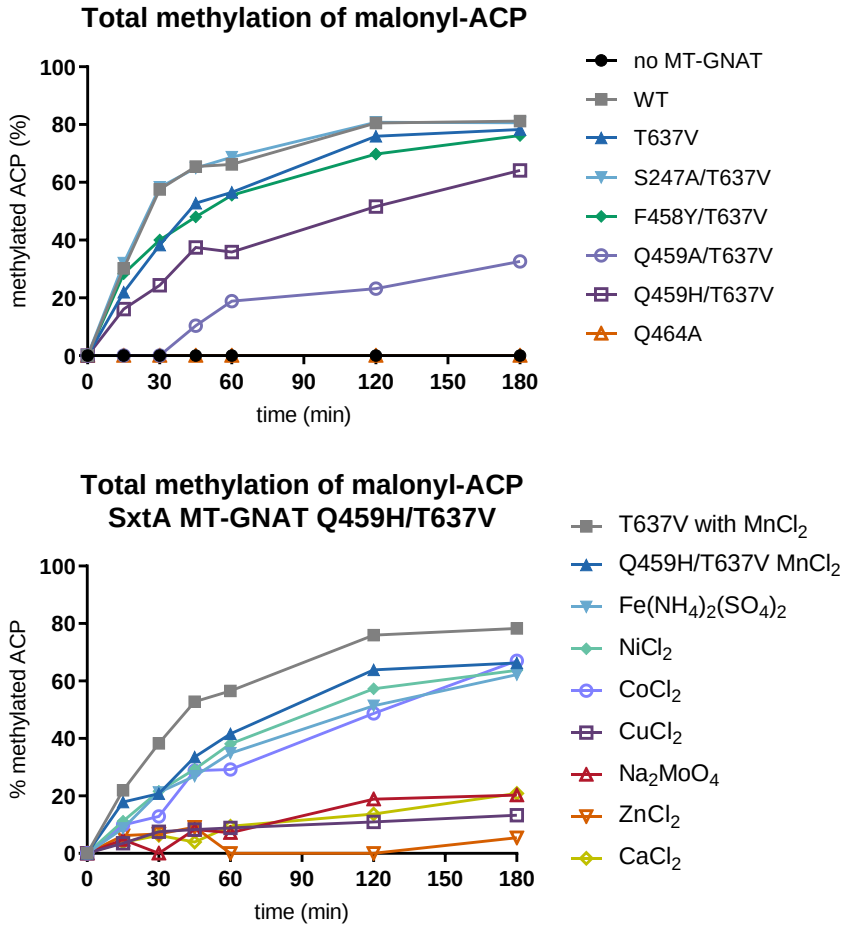


Figure 5.S3. Total methylation of malonyl-ACP, the sum of methylmalonyl- and propionyl-ACP population fractions, plotted per timepoint.
Mutations in metal-coordinating residues and deprotonation candidates (top) and metal salt screen with SxtA MT-GNAT Q459H/T637V.

UV-Vis spectroscopy of serine palmitoyltransferase (SPT): 20 µM of SPT and 0-10 mM of serine in 20 mM potassium phosphate pH 7.4 and 150 mM NaCl were combined for a total volume of 100 µL in a clear 96-well plate. The reactions equilibrated for 10 min at rt before spectroscopic

readings. Absorbance spectra were acquired from 300–550 nm, with the buffer only background spectra subtracted to obtain the traces in Figure 5.11.

UV-Vis spectroscopy of SxtA AONS: All reactions contained 20 µM of *M. wollei* SxtA AONS in 50 mM HEPES buffer with 150 mM NaCl for a total volume of 100 µL in a clear 96-well plate. The reactions equilibrated for 10 min at rt before spectroscopic readings. Absorbance spectra were acquired from 300–550 nm, with the buffer only background spectra subtracted to obtain the traces in Figure 5.10.

Arginine and arginine methyl ester titration: 0-10 mM at pH 7.0.

L-Arg-OMe at high pH: 4 mM methyl ester, with HEPES buffer at pH 7.5, 8.0, 8.2, 8.4 and 8.6. After initial readings, 1 mM propionyl-pantetheine (**5.32**) was added, and the spectra collected again after 10 min of equilibration.

*Titration of the native product (**5.7**):* 0-2 mM at pH 7.0.

*pH screen of ketone **5.7***: 1 mM of synthetic ketone **5.7** ranging from pH 6.6 to 8.6.

K_d of L-Arg and thiol/thioesters shifts in the absence of arginine (Figure 5.12): these data were collected by Sarah Ackenhusen. At pH 7.0, arginine was titrated in small increments, and peak changes at 326 and 425 nm were monitored to calculate the K_d. In Figure 5.12, the concentration of propionyl-pantetheine (**5.32**), propionyl-CoA and CoASH was 1 mM.

UV-Vis spectroscopy of arginine-based synthetic ketones: 20 µM of SxtA AONS and 1 mM ketones (as synthesized in Chapters 2 and 3) in 50 mM HEPES buffer, pH 7.0 or 8.0 and 150 mM NaCl were combined for a total volume of 100 µL in a clear 96-well plate. The reactions

equilibrated for 10 min at rt before spectroscopic readings. Absorbance spectra were acquired from 300–550 nm, with the buffer only background spectra subtracted to obtain the data in Figure 5.S4.

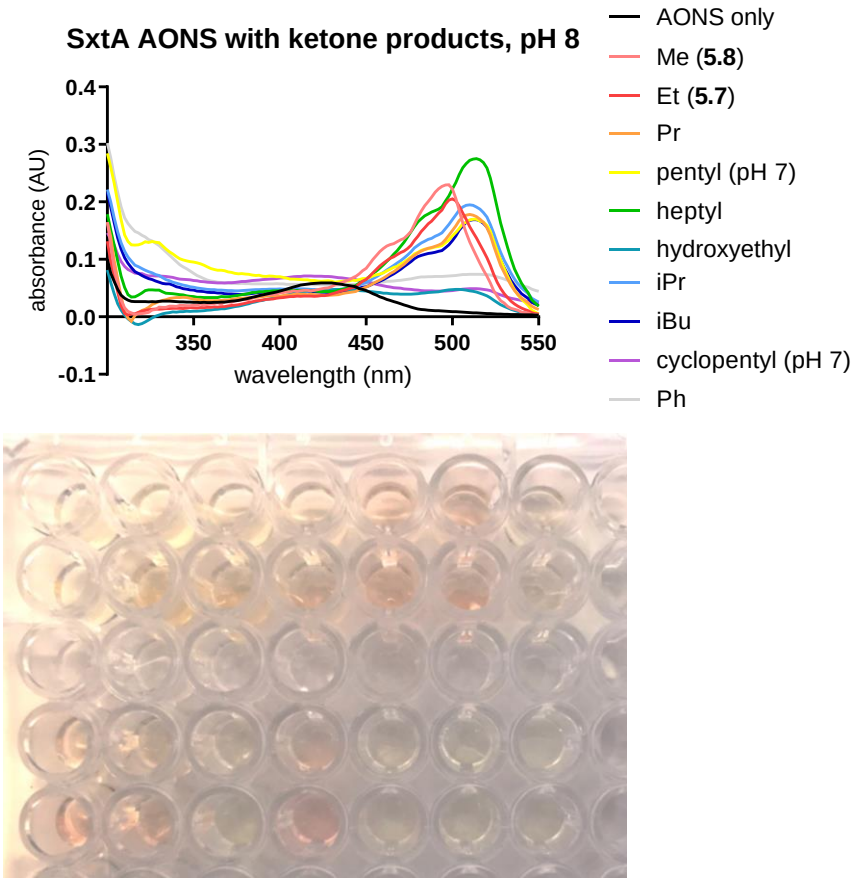
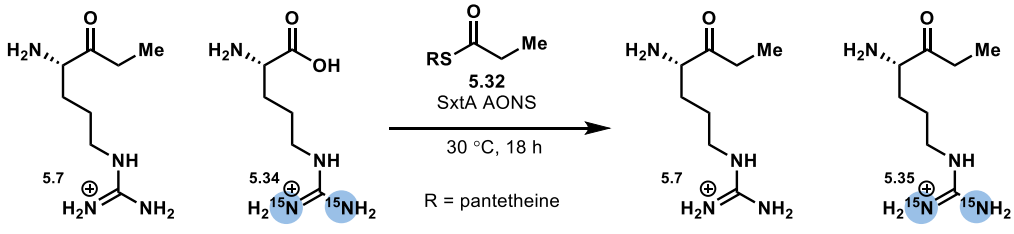


Figure 5.S4. SxtA AONS mixed with synthetic ketone product standards.
Top: UV-Vis absorbance spectra. Bottom: the ketone addition caused the enzymatic solutions to shift from pale yellow (bottom right) to orange and red (top rows). The addition of L-Arg to SxtA AONS caused solutions to turn clear (not shown).



Product inhibition of SxtA AONS: 2 mM of ^{15}N-labeled arginine (**5.34**, purchased from Cambridge Isotopes), 0-250 µM of synthetic ketone **5.7**, 1 mM propionyl-pantetheine (**5.32**) and

20 µM *M. wollei* SxtA ACP-AONS (see Chapter 4) in 50 mM HEPES pH 7.0, 150 NaCl were combined for a total reaction volume of 100 µL and incubated at 30 αC. For reactions with 0, 50 or 100 µM of synthetic product doped in, additional DMSO was added (2.5 2.0 or 1.5 µL, respectively) to maintain a constant volume of DMSO across all reactions. At 0.5, 1, 2, 3, 4, 5, 6 and 18 h, 10 µL of reaction was removed and quenched in 40 µL of methanol in a 96-well plate. Precipitation in the quenched reaction mixtures was pelleted by centrifugation of the plate at 2000 x *g* for 10 min and then filtering the supernatant through a 0.2 µM sterile filtration plate. 10 µL of the flow-through was added to a solution of 93 µL of acetonitrile, 95 µL of 1% v/v formic acid in water and 2 µL of 200 µM of tryptophan. Samples were analyzed by LC-MS: column = Waters Acquity 1.7 µm UPLC BEH Amide HILIC 2.1 x 100 mm column; method = 85% B at 0.3 mL/min for 2 min, followed by a linear gradient to 75% B at 0.4 mL/min over 3 min and a second linear gradient to 60% B over 0.5 min, 60% B for 1 min and then 6.5 min re-equilibration at 85% B at 0.3 mL/min (total time 15 min). t_R (**5.7**, **5.35**): 1.94 min; t_R (Trp): 1.55 min. The amount of labeled product formed was calculated by comparing the areas under the curve of the extracted ion chromatograms of **5.35** to the internal **5.7** standards.

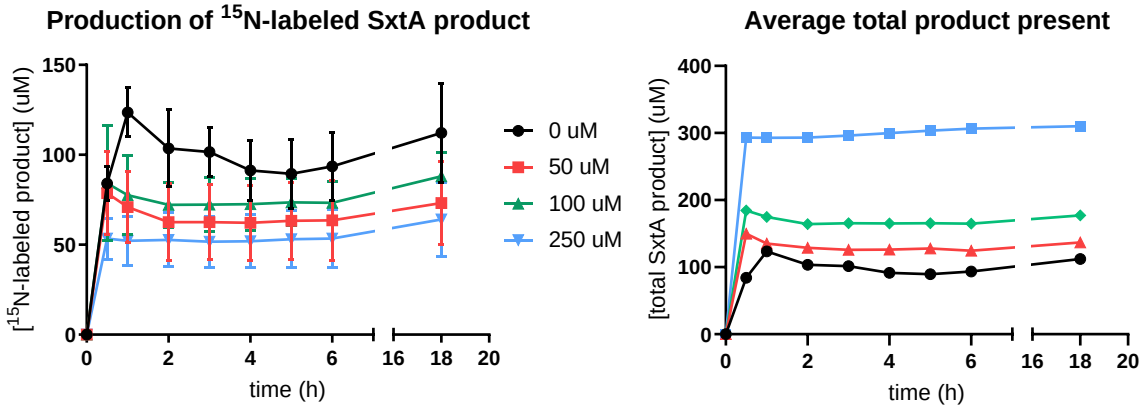


Figure 5.S5. Comparison of new ^{15}N-labeled ketone formed to the total product.
The error bars were not shown in Figure 5.13, but their large magnitude suggests inconsistent ketone synthesis, obscuring more detailed product inhibition analysis.

5.6 References

(1) Kozbial, P. Z.; Mushegian, A. R. Natural History of S-Adenosylmethionine-Binding Proteins. *BMC Struct. Biol.* **2005**, *5*, 19.

(2) Eliot, A. C.; Kirsch, J. F. Pyridoxal Phosphate Enzymes: Mechanistic, Structural, and Evolutionary Considerations. *Annu. Rev. Biochem.* **2004**, *73*, 383–415.

(3) Chun, S. W.; Hinze, M. E.; Skiba, M. A.; Narayan, A. R. H. Chemistry of a Unique Polyketide-like Synthase. *J. Am. Chem. Soc.* **2018**, *140*, 2430–2433.

(4) Skiba, M. A.; Sikkema, A. P.; Moss, N. A.; Tran, C. L.; Sturgis, R. M.; Gerwick, L.; Gerwick, W. H.; Sherman, D. H.; Smith, J. L. A Mononuclear Iron-Dependent Methyltransferase Catalyzes Initial Steps in Assembly of the Apratoxin A Polyketide Starter Unit. *ACS Chem. Biol.* **2017**, *12*, 3039–3048.

(5) Grindberg, R. V.; Ishoey, T.; Brinza, D.; Esquenazi, E.; Coates, R. C.; Liu, W.; Gerwick, L.; Dorrestein, P. C.; Pevzner, P.; Lasken, R.; Gerwick, W. H. Single Cell Genome Amplification Accelerates Identification of the Apratoxin Biosynthetic Pathway from a Complex Microbial Assemblage. *PLoS ONE* **2011**, *6*, e18565.

(6) Skiba, M. A.; Sikkema, A. P.; Fiers, W. D.; Gerwick, W. H.; Sherman, D. H.; Aldrich, C. C.; Smith, J. L. Domain Organization and Active Site Architecture of a Polyketide Synthase C-Methyltransferase. *ACS Chem. Biol.* **2016**, *11*, 3319–3327.

(7) Storm, P. A.; Herbst, D. A.; Maier, T.; Townsend, C. A. Functional and Structural Analysis of Programmed C-Methylation in the Biosynthesis of the Fungal Polyketide Citrinin. *Cell Chem. Biol.* **2017**, *24*, 316–325.

(8) Zou, X.-W.; Liu, Y.-C.; Hsu, N.-S.; Huang, C.-J.; Lyu, S.-Y.; Chan, H.-C.; Chang, C.-Y.; Yeh, H.-W.; Lin, K.-H.; Wu, C.-J.; Tsai, M.-D.; Li, T.-L. Structure and Mechanism of a Nonhaem-Iron SAM-Dependent C-Methyltransferase and Its Engineering to a Hydratase and an O-Methyltransferase. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2014**, *70*, 1549–1560.

(9) Kellmann, R.; Mihali, T. K.; Young, J. J.; Pickford, R.; Pomati, F.; Neilan, B. A. Biosynthetic Intermediate Analysis and Functional Homology Reveal a Saxitoxin Gene Cluster in Cyanobacteria. *Appl. Environ. Microbiol.* **2008**, *74*, 4044–4053.

(10) Young, J.; Stevens, D. C.; Carmichael, R.; Tan, J.; Rachid, S.; Boddy, C. N.; Müller, R.; Taylor, R. E. Elucidation of Gephyronic Acid Biosynthetic Pathway Revealed Unexpected SAM-Dependent Methylations. *J. Nat. Prod.* **2013**, *76*, 2269–2276.

(11) Poust, S.; Phelan, R. M.; Deng, K.; Katz, L.; Petzold, C. J.; Keasling, J. D. Divergent Mechanistic Routes for the Formation of Gem-Dimethyl Groups in the Biosynthesis of Complex Polyketides. *Angew. Chem. Int. Ed.* **2015**, *54*, 2370–2373.

(12) Limbach, H. H.; Chan-Huot, M.; Sharif, S.; Tolstoy, P. M.; Shenderovich, I. G.; Denisov, G. S.; Toney, M. D. Critical Hydrogen Bonds and Protonation States of Pyridoxal 5′-Phosphate Revealed by NMR. *Biochim. Biophys. Acta - Proteins Proteomics* **2011**, *1814*, 1426–1437.

(13) Raman, M. C. C.; Johnson, K. A.; Clarke, D. J.; Naismith, J. H.; Campopiano, D. J. The Serine Palmitoyltransferase from Sphingomonas Wittichii RW1: An Interesting Link to an Unusual Acyl Carrier Protein. *Biopolymers* **2010**, *93*, 811–822.

(14) Wadsworth, J. M.; Clarke, D. J.; Mcmahon, S. A.; Lowther, J. P.; Beattie, A. E.; Langridge-Smith, P. R. R.; Broughton, H. B.; Dunn, T. M.; Naismith, J. H.; Campopiano, D. J. The Chemical Basis of Serine Palmitoyltransferase Inhibition by Myriocin. *J. Am. Chem. Soc.* **2013**, *135*, 14276–14285.

(15) Ikushiro, H.; Fujii, S.; Shiraiwa, Y.; Hayashi, H. Acceleration of the Substrate Cα Deprotonation by an Analogue of the Second Substrate Palmitoyl-CoA in Serine Palmitoyltransferase. *J. Biol. Chem.* **2008**, *283*, 7542–7553.
(16) Jahan, N.; Potter, J. A.; Sheikh, M. A.; Botting, C. H.; Shirran, S. L.; Westwood, N. J.; Taylor, G. L. Insights into the Biosynthesis of the Vibrio Cholerae Major Autoinducer CAI-1 from the Crystal Structure of the PLP-Dependent Enzyme CqsA. *J. Mol. Biol.* **2009**, *392*, 763–773.
(17) Webster, S. P.; Alexeev, D.; Campopiano, D. J.; Watt, R. M.; Alexeeva, M.; Sawyer, L.; Baxter, R. L. Mechanism of 8-Amino-7-Oxononanoate Synthase: Spectroscopic, Kinetic, and Crystallographic Studies. *Biochemistry* **2000**, *39*, 516–528.
(18) Weeks, C. L.; Singh, S.; Madzelan, P.; Banerjee, R.; Spiro, T. G. Heme Regulation of Human Cystathionine β-Synthase Activity: Insights from Fluorescence and Raman Spectroscopy. *J. Am. Chem. Soc.* **2009**, *131*, 12809–12816.
(19) Kerbarh, O.; Campopiano, D. J.; Baxter, R. L. Mechanism of Alpha-Oxoamine Synthases: Identification of the Intermediate Claisen Product in the 8-Amino-7-Oxononanoate Synthase Reaction. *Chem. Commun.* **2006**, *18*, 60–62.
(20) Gong, J.; Hunter, G. A.; Ferreira, G. C. Aspartate-279 in Aminolevulinate Synthase Affects Enzyme Catalysis through Enhancing the Function of the Pyridoxal 5′-Phosphate Cofactor. *Biochemistry* **1998**, *2960*, 3509–3517.
(21) Pinon, V.; Ravanel, S.; Douce, R.; Alban, C. Biotin Synthesis in Plants. The First Committed Step of the Pathway Is Catalyzed by a Cytosolic 7-Keto-8-Aminopelargonic Acid Synthase. *Plant Physiol.* **2005**, *139*, 1666–1676.
(22) Bhor, V. M.; Dev, S.; Vasanthakumar, G. R.; Kumar, P.; Sinha, S.; Surolia, A. Broad Substrate Stereospecificity of the Mycobacterium Tuberculosis 7-Keto-8-Aminopelargonic Acid Synthase: Spectroscopic and Kinetic Studies. *J. Biol. Chem.* **2006**, *281*, 25076–25088.
(23) Hunter, G. A.; Ferreira, G. C. Pre-Steady-State Reaction of 5-Aminolevulinate Synthase. *J. Biol. Chem.* **1999**, *274*, 12222–12228.
(24) Turbeville, T. D.; Zhang, J.; Christopher Adams, W.; Hunter, G. A.; Ferreira, G. C. Functional Asymmetry for the Active Sites of Linked 5-Aminolevulinate Synthase and 8-Amino-7-Oxononanoate Synthase. *Arch. Biochem. Biophys.* **2011**, *511*, 107–117.
(25) Eschenfeldt, W. H.; Lucy, S.; Millard, C. S.; Joachimiak, A.; Mark, I. D. A Family of LIC Vectors for High-Throughput Cloning and Purification of Proteins. In *Methods in Molecular Biology*; Doyle, S. A., Ed.; Humana Press: Totowa, NJ, 2009; Vol. 498, pp 105–115.
(26) Dorrestein, P. C.; Bumpus, S. B.; Calderone, C. T.; Garneau-Tsodikova, S.; Aron, Z. D.; Straight, P. D.; Kolter, R.; Walsh, C. T.; Kelleher, N. L. Facile Detection of Acyl and Peptidyl Intermediates on Thiotemplate Carrier Domains via Phosphopantetheinyl Elimination Reactions during Tandem Mass Spectrometry. *Biochemistry* **2006**, *45*, 12756–12766.

Chapter 6: Synthesis of chiral amines with type II oxidase AnaB

Summary

Nearly 40% of bioactive pharmaceutical compounds contain a chiral amine motif, but efficient chemo- and stereoselective preparations of these compounds remains challenging. Extensively functionalized amines, including the cyclic amines, are found in NRPS pathways but are also relatively expensive to utilize in chemoenzymatic reactions. We identified a highly reactive CP-bound intermediate generated by the oxidase AnaB in anatoxin-a biosynthesis that could be harnessed for the preparation of chiral cyclic amines. The released small molecule products could serve as a building block to more complex bioactive scaffolds but cannot be efficiently or economically generated with stoichiometric carrier protein AnaD. Here, a four-step carrier protein-partner enzyme catalytic cycle is developed to load an amine substrate, oxidize to an iminium ion by AnaB, add an exogenous nucleophile stereoselectively and offload product in a self-contained cycle that allows for the use of catalytic quantities of CP. Although we identified complementary loading and offloading cycles and a nucleophile and confirmed AnaB activity, the complete one-pot cycle is stalled by the formation of an overoxidized side product. Future development of this platform will be aimed at limiting the undesired overoxidation and optimizing preparative-scale chemoenzymatic reactions to isolate small molecule products in synthetically relevant quantities.

6.1 Introduction

Although we initially selected the polyketide-like synthase SxtA to develop a seemingly self-contained unit with its own loading and offloading, our proposed carrier protein-partner enzyme platforms are not limited to type I *cis*-acting systems. We have also investigated standalone enzymes from another secondary metabolite biosynthetic pathway to develop catalytic cycles where all enzymes are acting in *trans*.

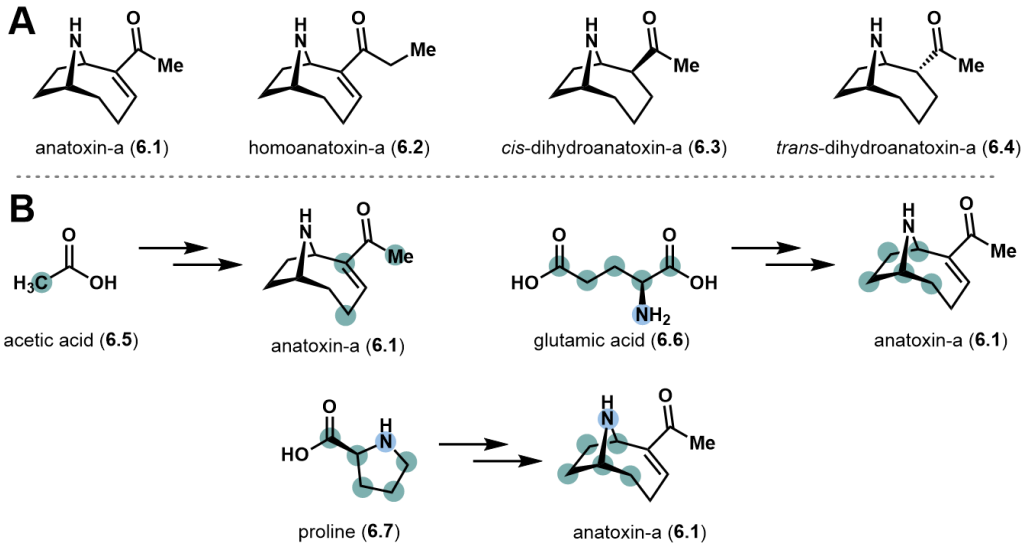


Figure 6.1. The natural product anatoxin-a.
(A) The parent toxin and some common derivatives. (B) Feeding studies with anatoxin-a producers by Hemscheidt and Ploux.

Anatoxin-a (**6.1**) and naturally produced analogs (**6.2-6.4**) are potent toxins produced by several genera of cyanobacteria, including *Anabaena flos-aquae* (Figure 6.1A). Like saxitoxin, it is a component of algal blooms, and first discovered after poisoning cattle and dogs in Canada who drank contaminated water in the 1960s.[1] The toxin was first isolated in 1977 and reported with to possess a bicyclic structure.[2] Aronstam and Witkop then reported in 1981 that anatoxin-a binds irreversibly to neuronal- and muscle-type nicotinic acetylcholine receptors (nAChRs).[3] With toxin bound to nAChRs, the associated sodium channels initially open to release Na^+ ions, and then eventually desensitize to block additional ion flow, resulting in respiratory paralysis and death.[3]

The similarly named molecule anatoxin-a(S) (now known as guanitoxin) has the same mechanism of action,[4] but the two compounds are structurally unrelated. Anatoxin-a is much less toxic than saxitoxin (LD_{50} of 200-250 µg/kg vs. approx. 8.3 ng/kg of STX)[5,6] by intraperitoneal injection in mice, but a toxic dose of anatoxin-a is fatal in under 30 min, earning it the nickname very fast death factor.[1]

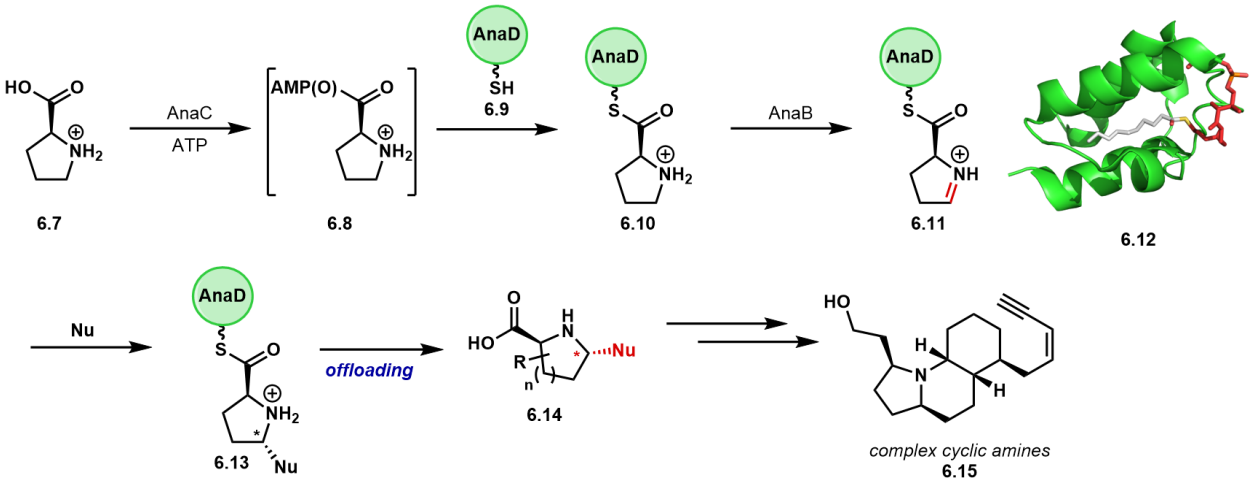


Figure 6.2. Initiation of anatoxin-a biosynthesis and oxidation Pro-AnaD by the flavin-dependent oxidase AnaB and elaboration into a chiral amine with a second stereocenter. ACP **6.12** is an acyl carrier protein with a fatty acyl chain bound and protected with the three α-helices. PDB: 2FAE.

From feeding experiments of labeled acetic acid and glutamic acid in the mid-1990s, Hemscheidt and coworkers proposed that the bridging bicyclic core of anatoxin-a was biosynthesized from glutamic acid (**6.6**), proline or ornithine and the remainder of the molecule by a polyketide synthase (Figure 6.1B).[7] By 2009, Ploux and coworkers had identified a polyketide synthase and nonribosomal peptide synthetase biosynthetic gene cluster unique to anatoxin-a-producing strains, named *ana*.[8] Further feeding studies indicated that the starter unit of the anatoxin-a biosynthetic pathway is L-proline (**6.7**), like other pyrrolidine- or pyrrole-containing natural products.[9,10] The Ploux group has demonstrated that L-Pro is adenylated by AnaC, and loaded onto the peptidyl carrier protein (PCP) AnaD (see **6.10**, Figure 6.2).[10] The subsequent two-

electron oxidation of **6.10** is performed by the flavin-adenine dinucleotide (FAD)-dependent oxidase AnaB.[11] The product, iminium ion **6.11**, contains an electrophilic carbon center that we envision leveraging for the synthesis of chiral cyclic amines in an artificial CP-PE catalytic cycle. The iminium ion species **6.11** is highly reactive and could potentially be functionalized by a nucleophile. In addition, because CPs may protect reactive cargo in a chiral environment within their helices (see **6.12**),[12] the nucleophilic addition could be stereoselective (**6.14**), forming a cyclic amine with two stereocenters.

Chiral amines are a ubiquitous motif in pharmaceutical compounds, represented in about 40% of products on the market in 2014.[13] Like the preparation and modification of amino acids, chiral amine synthesis often faces challenges in protecting group manipulation, stereoselectivity, and chiral resolution, among others.[14] We envisioned developing the AnaD-AnaB pair to oxidize proline-related heterocycles and the subsequent nucleophilic addition step as a biocatalytic platform to synthesize chiral amines. However, this would require building loading and offloading steps into the total catalytic cycle; two potential routes with differing loading methods are outlined in Figure 6.3.

The *apo* catalytic cycle (Figure 6.3A) starts with the inactive *apo* form of AnaD. A promiscuous phosphopantetheinyl transferase (PPTase) transfers both Ppant and substrate to *apo*-AnaD (**6.16**) from coenzyme A (CoA) thioesters. The substrate may then be oxidized up to the iminium ion **6.11** and attacked by a nucleophile to form **6.13**. Acyl carrier protein hydrolase (AcpH) closes the cycle by cleaving off the Ppant-product thioester **6.19**. To reduce the cost of forming Pro-CoA (**6.17**), we also can start from a precursor pantetheine (Pant) thioester **6.18** and build up to CoA through the CoA biosynthetic enzymes CoaADE.[15] Alternatively, in a catalytic cycle starting from *holo*-AnaD (**6.9**, Figure 6.3B), activated substrates are loaded diffusively by

nonenzymatic transthioesterification.[16] Following oxidation by AnaB and nucleophilic addition, the product acid **6.14** is cleaved from AnaD by a type II thioesterase (TEII). Steps were verified separately before successively combination into the full four-step process.

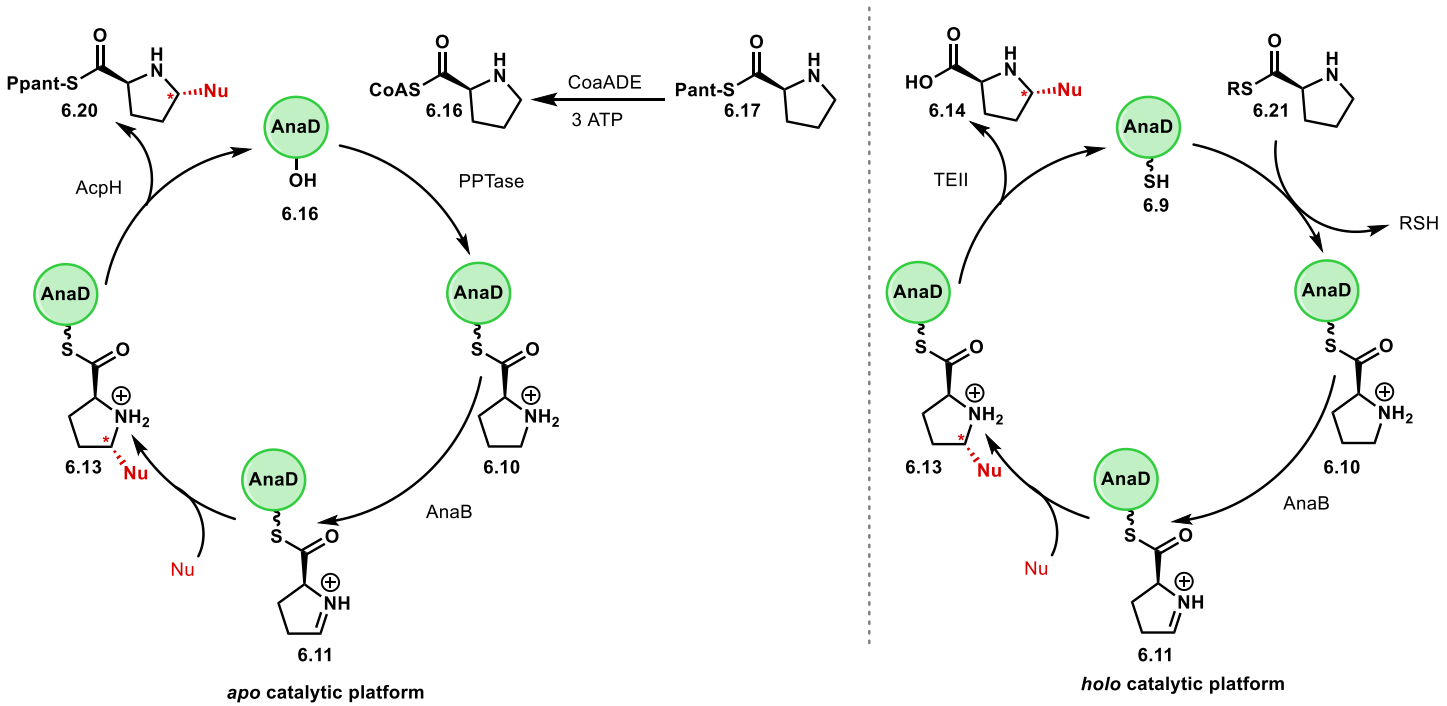


Figure 6.3. Generating significant quantities of chiral amines by AnaB and nucleophiles also requires the development of a platform to utilize the carrier protein AnaD in catalytic amounts as well.
Potential catalytic cycles can begin with either *apo*- or *holo*-AnaD. PPTase: phosphopantetheinyl transferase; TEII: type II thioesterase; ATP: adenosine triphosphate.

6.2 Loading and Offloading

From Apo-AnaD

The two proposed catalytic cycles use two different paired loading and offloading methods. For the *apo* cycle (Figure 6.3A), both steps were envisioned to be enzymatic. We used the promiscuous PPTase Sfp from *Bacillus subtilis*, which natively primes *apo*- to *holo*-PCPs in surfactin biosynthesism as the catalyst for the loading step.[17] CoaADE, two kinases (CoaAE) and one adenylyltransferase (CoaD) are three enzymes from coenzyme A biosynthesis, and have been shown to transform an extremely wide range of pantetheinyl derivatives into CoA compounds using 3 ATP molecules.[15,18] We synthesized Pro and phenylalanine aminoacyl-CoA thioesters as the substrate for Sfp, and two precursors, Pro- and Phe-pantetheine (see **6.21** and **6.22**, Figure 6.4A) for CoA biosynthesis.

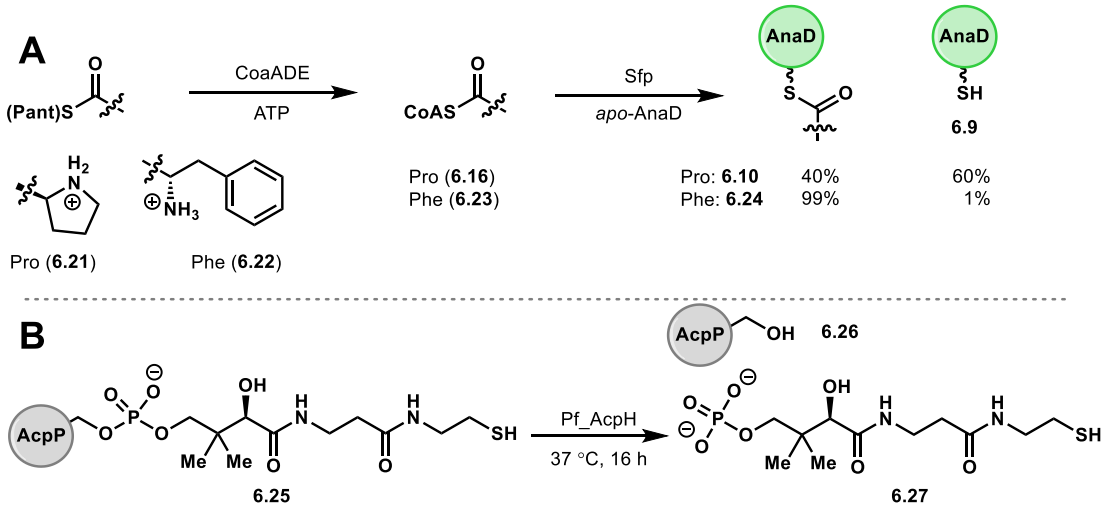


Figure 6.4. Enzymatic loading and offloading reactions for the *apo* catalytic cycle.
(A) CoA biosynthesis from pantetheine thioesters, followed by priming/loading catalyzed by Sfp. (B) AcpH from *P. fluorescens* converts its native substrate *holo*-AcpP to *apo* but is inactive with AnaD.

Plasmids containing *anaD* and *anaB* from the anatoxin-a producing strain *Oscillatoria* PCC 6506 were obtained from Olivier Ploux at the French National Centre for Scientific Research.[9,10]

Although AnaD expressed well in BL21(DE3) and BAP1 strains of *E. coli*,[19] the *apo*-PCP precipitated in the low salt storage buffer, rendering quantification difficult. Fortunately, the precipitated protein re-solubilized in enzymatic reactions to prime and load substrates. By contrast, *holo*-AnaD was soluble in all conditions used.

Incubation of *apo*-AnaD and Sfp with synthetic and biosynthetically produced Pro-CoA (**6.16**) or Phe-CoA (**6.23**) and analysis by intact protein mass spectrometry showed the expected mass of aminoacyl-AnaD (**6.10** and **6.24**, Figure 6.4A). For Pro, the major product was *holo*-AnaD, an observation also noted by the Ploux lab.[10] We propose that the hydrolysis of Pro is primarily an artifact of ionization, as proline and other secondary amines are known to be especially reactive in the gas phase.[20]

To regenerate *apo*-AnaD, we obtained the genes for AcpH from *Pseudomonas aeruginosa* and *Psueodomonas fluorescens*.[21–23] Endogenously, AcpH is a phosphodiesterase that cleaves the Ppant cofactor from *holo*-AcpP, the acyl carrier protein of fatty acid biosynthesis in primary metabolism, to *apo*-AcpP (Figure 6.4B).[21] This apparently nonessential protein has been reported to aggregate easily,[21,23] but active AcpH has also been demonstrated to operate on AcpP with substrates bound and non-native CPs.[18] *P. fluorescens* is reported to be both the more stable and promiscuous homolog.[23] We first expressed the two AcpH homologs from *P. aeruginosa* (Pa) and *P. fluorescenes* (Pf). The former was inactive in our hands, while Pf_AcpH was also not active when expressed under the T7 promoter in DE3 strains. With assistance of Dr. Gregory Dodge of Prof. Janet Smith's group at the University of Michigan, *Pf_acpH* was placed under control of the Trc[24] promoter. The new recombinant protein successfully converted *E. coli holo*-AcpP completely to apoprotein and Ppant (**6.27**) but had no activity on *holo*-AnaD. Sequence alignment of AcpP and AnaD shows only 36% identity between the two proteins; it was not unexpected that

Pf_AcpH and AnaD are incompatible. Considering the difficulty in obtaining active Pf_AcpH, we then turned our attention to the *holo* catalytic cycle, which also uses catalytic amounts of Ppant.

From Holo-AnaD

The second catalytic cycle starts from *holo*-CP and utilizes a paired offloading mechanism to return to the same species. We propose a nonenzymatic loading step to bypass the native loading gatekeeper AnaC[10] and additionally, eliminate the use of ATP. However, our aminoacyl donor substrates still require activation to thioesters before undergoing nonenzymatic transthioesterification with the Ppant cofactor of *holo*-AnaD. For this purpose, we synthesized the thiophenol (-SPh) and *N*-acetylcysteamine (-SNAC) thioesters of L-Pro and L-Phe (Figure 6.5). The thiophenol esters oxidized the *holo*-AnaD, forming a disulfide adduct (**6.30**, +108 Da) rather than aminoacyl-AnaD species. The SNAC thioesters afforded the expected aminoacyl-AnaD after incubation at 28 °C for 1-3 h. We have determined that the resulting free HSNAC causes significant amounts of hydrolysis back to *holo*-AnaD, so 2-vinylpyridine was added to scavenge the thiol.[16] Due to the structural similarity of HSNAC and Ppant, the scavenger forms an adduct with AnaD as well (**6.34**), and was not used in later experiments.

This second catalytic cycle regenerates *holo*-AnaD by TEIIs. Unlike type I thioesters that cleave finished PKS/NRPS products at the end of a pathway, the type II enzymes cleave stalled unproductive intermediates from CPs to restart biosynthesis.[25] Four TEIIs from NRPS pathways were expressed and purified, including AnaA, the putative TEII of the anatoxin-a pathway.[8,26–28] Pro- and Phe-loaded AnaD (by any of the above methods) was incubated with TEII for 2 h at 28 °C (Table 6.1). Although all four TEIIs showed some activity, TycF was generally the most active for both substrates and was chosen for further studies. SNAC thioesters are also hydrolyzed,

potentially diminishing the overall yield by prematurely cleaving and deactivating the aminoacyl donor before loading onto AnaD, but the Walsh lab has reported that CP-bound thioesters are

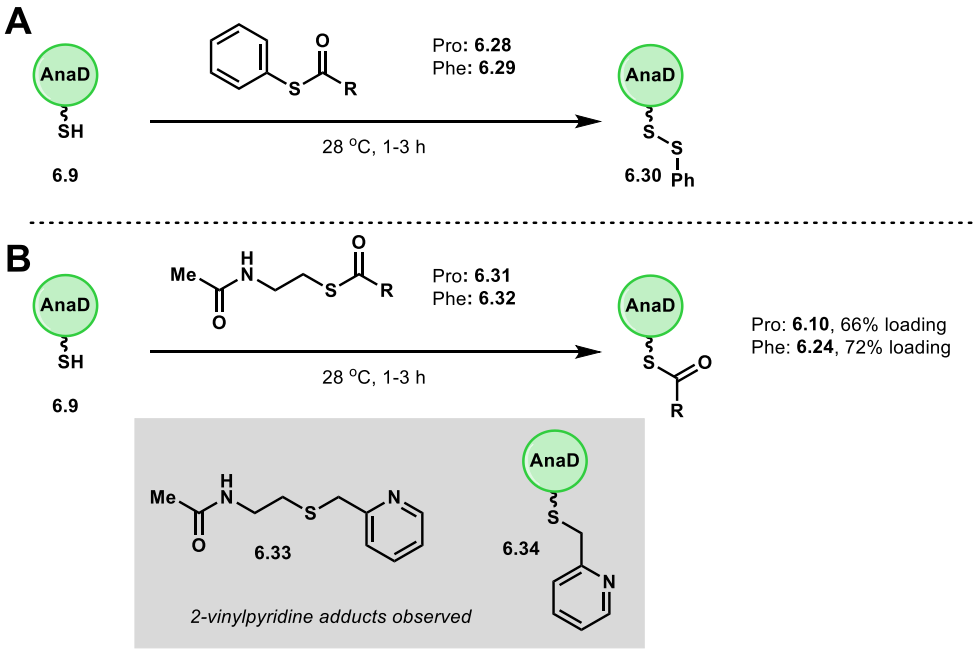
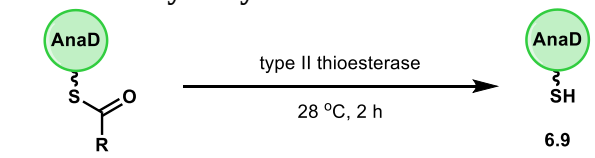


Figure 6.5. Diffusive loading of *holo*-AnaD by thiophenol and SNAC esters.
(A) Thiophenol esters oxidize the carrier protein itself. No aminoacyl-AnaD was observed. (B) SNAC thioesters transfer aminoacyl groups for incomplete loading. Adding 2-vinylpyridine as a thiol scavenger also inactivated *holo*-ACP.

favored over SNAC thioesters.[26] Activity toward all substrates appears to slow dramatically after 2 h, for unknown reasons, never reaching complete hydrolysis. Three additional genes of TEIIs (*grsT*, *pltG* and *srfAD*) from NRPS pathways have been obtained and will be expressed and assessed for activity.[29–31] With the loading and offloading steps for the *holo* catalytic cycle established, we then began combining steps of our artificial catalytic cycle in one pot.

Table 6.1. Hydrolysis of Pro and Phe aminoacyl groups from AnaD



Pro: **6.10**, 52% loading
Phe: **6.24**, 70% loading

Entry	Thioesterase	Fraction of *holo*-AnaD (**6.9**)	
		From **6.10** (Pro)	From **6.24** (Phe)
1	None (0 h)	48	30
2	None (2 h)	70	37
3	AnaA	76	50
4	TycF	78	79
5	RedJ	82	61
6	RifR	78	50

6.3 Multi-Step Catalytic Cycles

Of the known flavin-dependent Pro-PCP oxidases and dehydrogenases, AnaB is the only enzyme that performs a two-electron oxidation, while PltE,[32] PigA,[33] CloN5 and CouN5[34] catalyze four-electron oxidations to aromatic pyrrole species which are planar and no longer highly electrophilic. Ploux and coworkers have proposed a mechanism for AnaB-catalyzed oxidation of Pro-AnaD **6.10**, beginning with the α-deprotonation by Glu244, followed by hydride transfer to the FAD cofactor (Figure 6.6, top). A Glu244-assisted rearrangement affords the product with the unsaturation between N1 and C5. The reduced AnaB returns to the oxidized state with molecular oxygen, releasing a molecule of hydrogen peroxide.[11] Additionally, the iminium ion is proposed to be maintained throughout extension by PKS modules AnaEF until cyclization by AnaJ (Figure 6.6, bottom).[35]

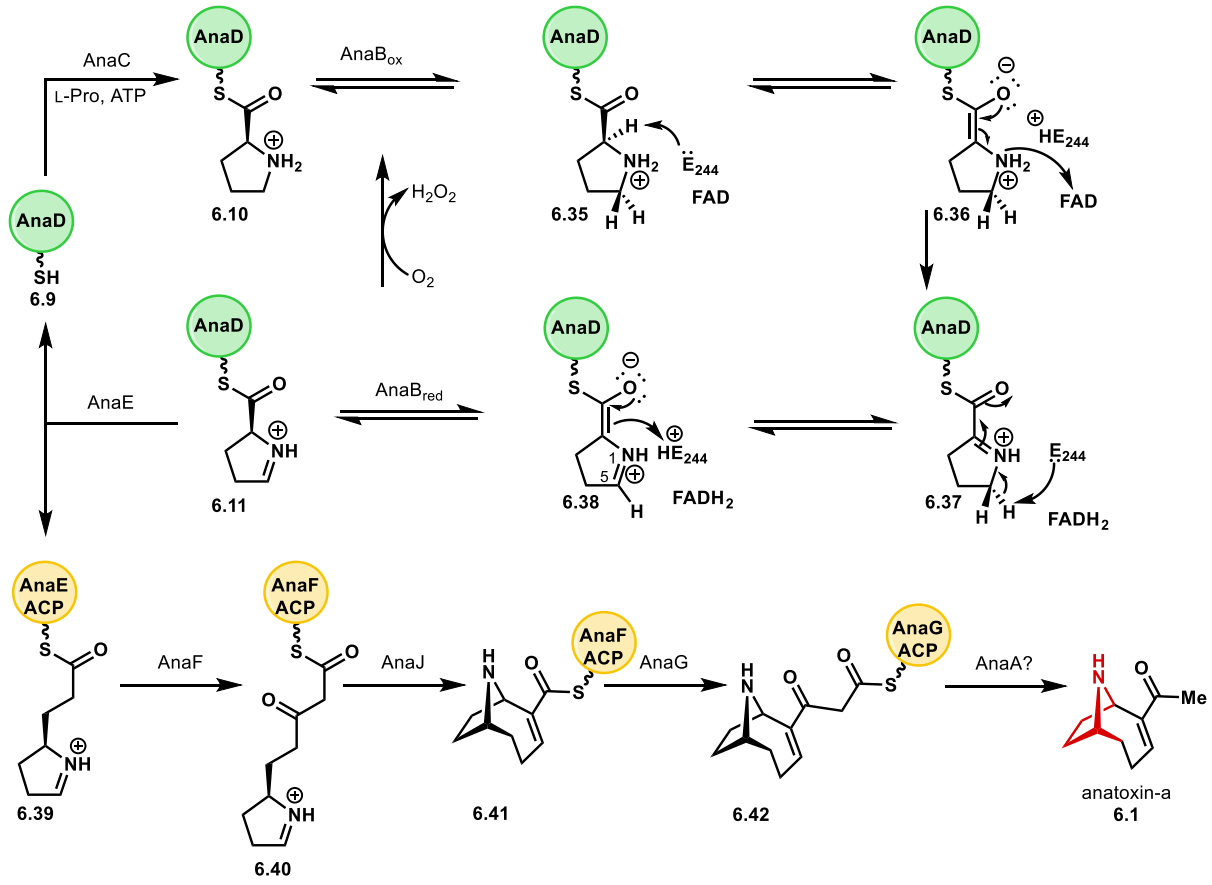


Figure 6.6. Proposed mechanism of AnaB to generate the first iminium intermediate, then elongation and cyclization into the final product anatoxin-a.

AnaB was incubated with free proline and proline thioesters (pantetheine, CoA, SNAC), but no oxidation was observed by MS. When AnaD-bound substrate (**6.10**) was treated with AnaB, the mass of **6.10** decreased by 2 Da after a short incubation with AnaB, corresponding to the structure of **6.11** (Figure 6.7). A 4 Da loss was also seen at extended times, indicating that a second undesired oxidation to pyrrole **6.43** is occurring. It is not yet clear whether this overoxidation is enzymatic or spontaneous but could be examined in an O_2-free single turnover oxidation reaction consisting of Pro-AnaD **6.10** and AnaB. Both oxidation products were confirmed by the MS/MS Ppant ejection assay.[36]

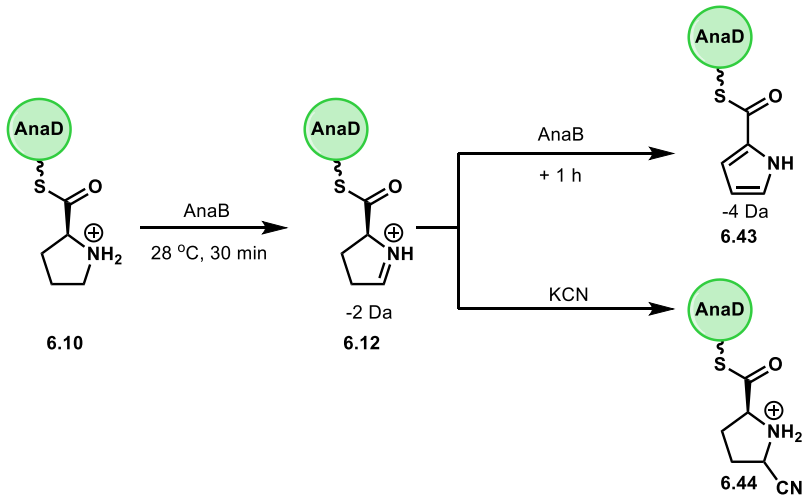
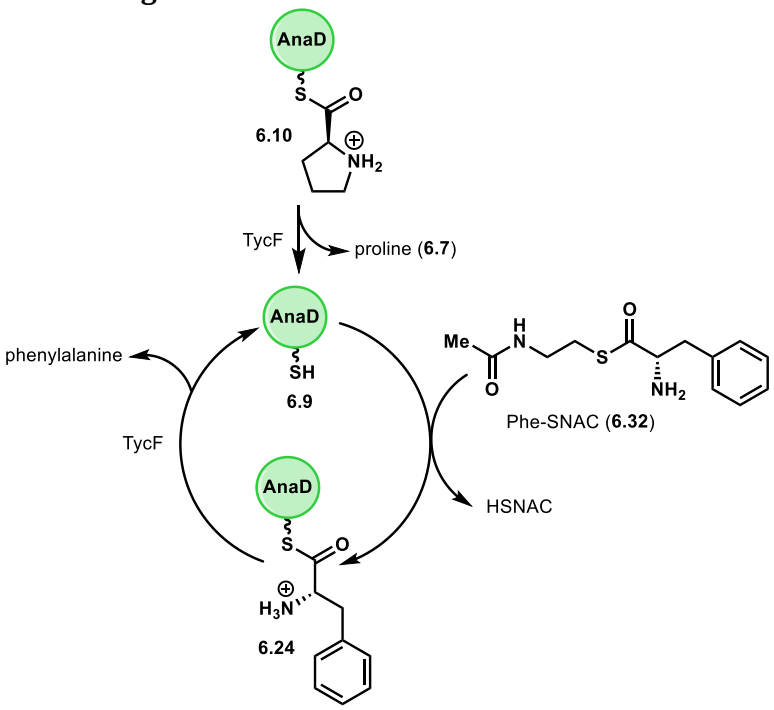


Figure 6.7. AnaB-catalyzed oxidation of Pro-AnaD to pyrrolinium and pyrrole species. The initial AnaB product (**6.12**) could be intercepted by nucleophilic addition before the second oxidation reaction that erodes all stereoinformation and renders the substrates inactive.

KCN, indole or nitromethane were also added after loading to act as a nucleophile, attacking intermediate **6.11** before the second oxidation could occur. Only a small amount of a species corresponding to cyanide addition (**6.44**) was detected.

With all four of the processes successful operating in isolation, two steps (loading and offloading) were then combined in one pot (Table 6.2) to demonstrate substrate flux throughout the system. AnaD pre-loaded with Pro (**6.10**) was incubated with TycF to offload AnaD, and Phe-SNAC (**6.32**) to load a second substrate (Table 6.2). Background hydrolysis to *holo*-AnaD (**6.9**) by the buffer or HSNAC occurred in all runs, as measured by intact protein MS (entry 2). Because of incomplete Pro loading, some Phe-loaded AnaD (**6.24**) was observed in the absence of TEII (entry 4). In the presence of both TycF and thioester **6.32**, there was increased conversion of **6.10** to *holo*-AnaD but less Phe loading, likely due to additional TycF-catalyzed hydrolysis (entry 5). The same pattern of reactivity was observed when pre-loaded **6.24** was incubated with TycF and Pro-SNAC (**6.32**).

Table 6.2. Substrate exchange from Pro- to Phe-AnaD



Entry	Time	TycF	Phe-SNAC	Pro (6.10)	*Holo* (6.9)	Phe (6.24)
1	0 min	-	-	52%	48%	0%
2	30 min	-	-	47	53	0
3	30 min	+	-	34	66	0
4	30 min	-	+	32	28	40
5	30 min	+	+	27	39	34

After confirming that loading and offloading were possible in one pot, three consecutive steps (loading, oxidation and offloading) were investigated (Figure 6.8). Small molecule analysis by LC-MS showed hydrolyzed proline (**6.7**) and the growth of a peak corresponding to the mass of acid **6.45**, but not hydroxylated proline **6.46** or acid **6.47**. Analysis of the intact proteins showed the accumulation of side product **6.43**, which we attribute to TycF having limited activity towards pyrroles bound to AnaD. Varying the pH or amounts of TycF or AnaB did not significantly improve the relative yield of the desired product **6.45**.

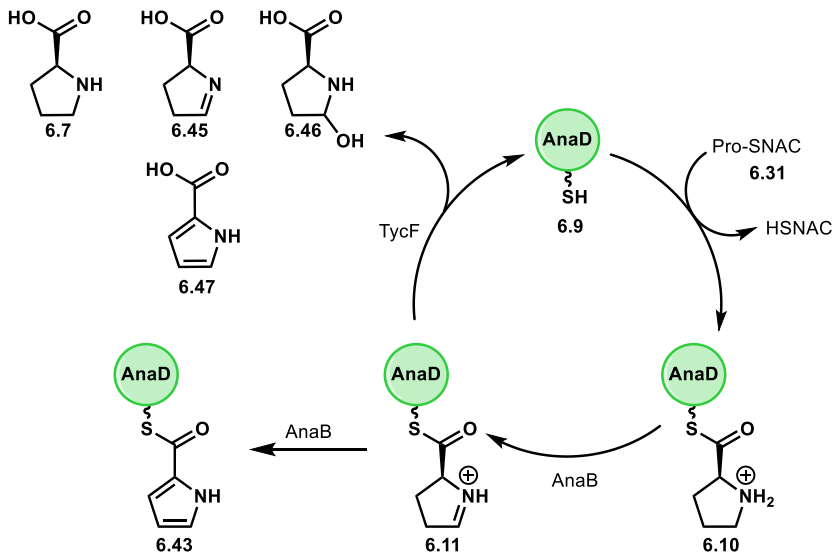


Figure 6.8. The basic three-step catalytic cycle consisting of substrate loading, oxidation by AnaB and product cleavage by TycF.

A complete one-pot four-step cycle (Figure 6.19) containing AnaD, AnaB, aminoacyl donor Pro-SNAC and the nucleophile KCN was incubated. By intact protein MS, a new peak and ejected fragment corresponding to **6.44** appeared after 1 h, but also diminished over extended time. When the small molecules were analyzed we saw none of the desired acid **6.47**, only **6.45** from premature cleavage. The nitrile Pro derivatives **6.44** and **6.47** may be unstable, so we can only confirm three of four processes currently. Although the side product **6.43** accumulated more slowly here than in the three-step cycle, we propose that the TycF-catalyzed cleavage to close the catalytic cycles is the main bottleneck to AnaD turnover. We will continue to study the kinetics of our catalytic cycle with new nucleophiles and TEIIs, and the effects of adjusting the amount of each component.

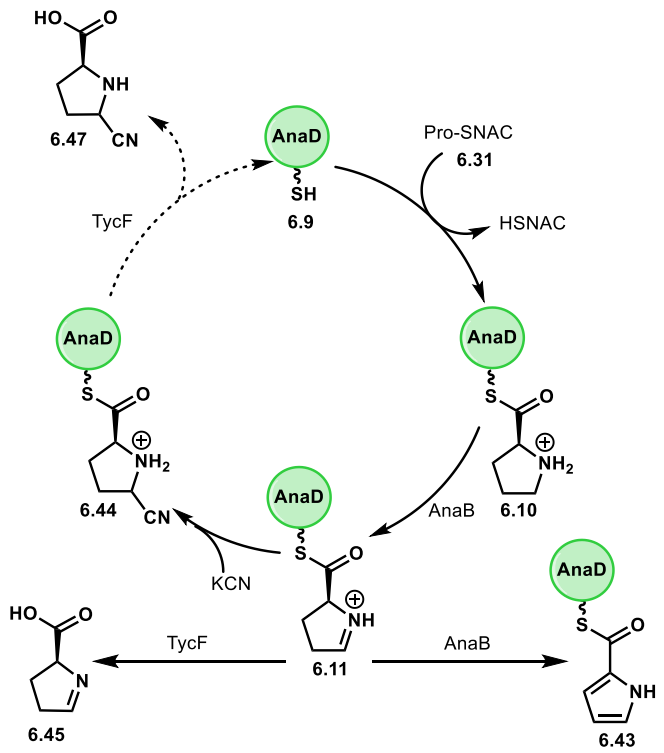


Figure 6.9. The preliminary four-step *holo* catalytic cycle in one pot to load Pro onto AnaD, AnaB-catalyzed oxidation, cyanide addition and cleavage by TycF only completed three of four steps.
Overoxidation to the pyrrole species **6.43** was still faster than nucleophilic addition.

Because the CP-PE platform also provides a method to evaluate AnaB's substrate scope, we have begun screening some unnatural substrates (Figure 6.10). Increasing the ring size to 6 (**6.48**) abolished AnaB activity. Unexpectedly, D-Pro (D-**6.10**), was oxidized in the three-step catalytic cycle, and a mass corresponding to **6.11** appeared, but with unknown configuration at C2. The crystal structure of AnaB shows the possible catalytically active His125 residue opposite the known active Glu244,[37] so mutagenesis efforts are in progress to generate E244A and H125A variants to determine their catalytic importance and identify other relevant residues.

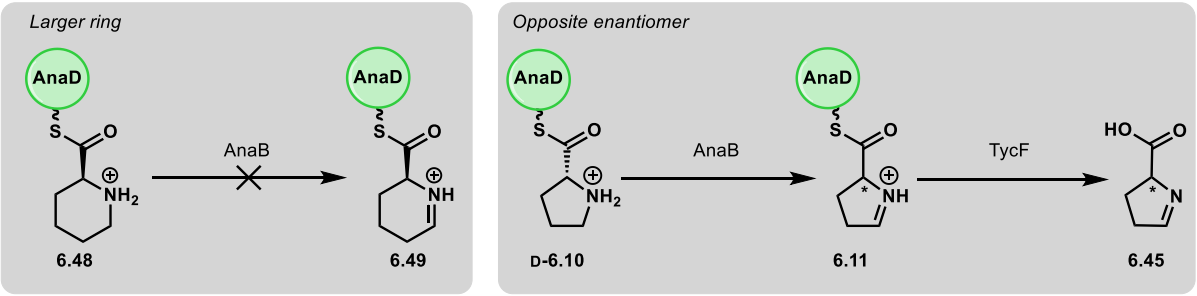


Figure 6.10. Reactivity of AnaB with non-native substrates.

6.4 Conclusions

Despite their ubiquity in pharmaceutical and fine chemical products, chiral cyclic amines present a major synthetic challenge. The biosynthesis of natural products in Nature provides the bases and inspiration for the development of new chemoenzymatic strategies. We designed a biocatalytic platform centered on the oxidation of a prolyl-AnaD species by the partner enzyme AnaB to the iminium intermediate **6.11**, that then can be attacked by exogenous nucleophiles to generate a second chiral center. In order to use AnaD catalytically, we also added loading and offloading steps to recycle the PCP. In the first catalytic cycle, *apo*-AnaD could be primed and loaded, but we could not regenerate the *apo* state with a phosphodiesterase originally intended for fatty acid synthase carrier proteins. In the second catalytic cycle, *holo*-AnaD was diffusively loaded by synthetic acyl donors and potential products were hydrolyzed by thioesterases, successfully closing the catalytic cycle. In addition, we observed AnaB-catalyzed oxidation and small amounts of cyanide addition to the iminium ion product.

However, we are not able to perform the complete four-step cycle in the same pot efficiently. This is primarily due to the overoxidation of **6.11** to a pyrrole species (**6.43**), which is stable and can no longer be attacked by cyanide or other water-compatible electrophiles, or even cleaved to

restart the catalytic cycle (Figure 6.11). Thus, our process is stalled by the pyrrole side product and future efforts to develop AnaB as a viable biocatalyst must focus on preventing formation of **6.43**.

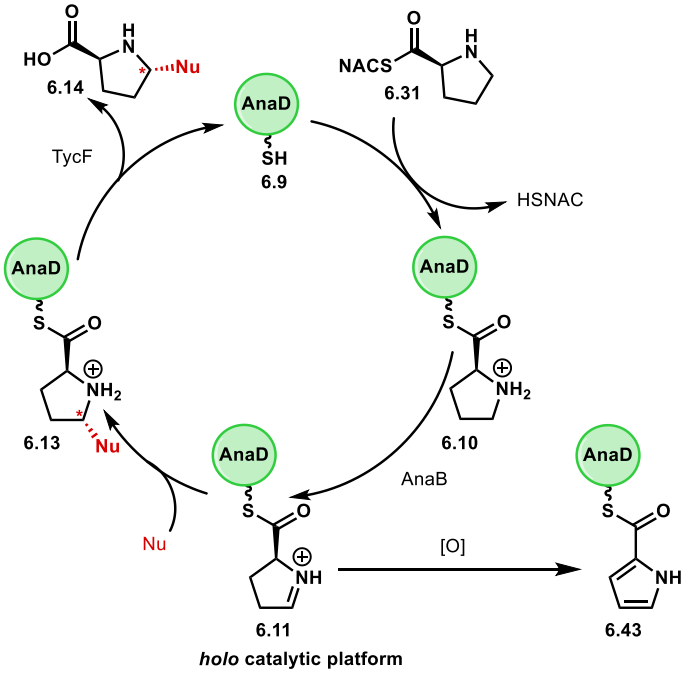


Figure 6.11. The final *holo* catalytic platform built in this Chapter.

Additional AnaB homologs from other anatoxin-a-producers can be tested to identify a protein that mediates the second oxidation reaction more slowly (if an enzymatic process). Regardless of whether the overoxidation is AnaB-catalyzed or spontaneous, more nucleophiles will be screened to discover a reagent that can trap the iminium intermediate more rapidly. Once the complete four-step catalytic cycle (Figure 6.11) has been demonstrated on the native Pro substrate, we will turn our attention to increasing the overall catalytic efficiency. This will likely focus on tuning the enzymatic activity of AnaB and TycF or other TEIIs by directed evolution.

The strategies investigated here to efficently utilize a PE that is absolutely CP-dependent (*in situ* substrate generation, nonenzymatic loading from inexpensive aminoacyl donors in combination with multiple offloading methods) provide preliminary groundwork for the

development of other artificial CP-PE biocatalytic platforms. The products of these catalytic cycles will not just be restricted to amino acid-based modifications or amines in NRPS systems, but also to other synthetically valuable scaffolds found in PKS and hybrid systems. With screening of many other CP-PE combinations and directed evolution, we anticipate that increasingly more PEs will soon be leveraged for their synthetic potential.

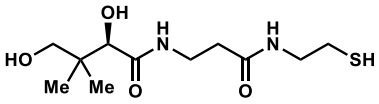
6.5 Experimental

I. Chemical synthesis

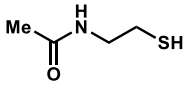
A. General information

All reagents were used as received unless otherwise noted. Reactions were carried out under a nitrogen atmosphere using standard Schlenck techniques unless otherwise noted. Solvents were degassed and dried over aluminum columns on an MBraun solvent system (Inert Corporation, Model PS-00-3). Reactions were monitored by thin layer chromatography using Machery-Nagel 60 F254 precoated silica TLC plates (0.25 mm) or Merck Silica Gel 60 RP-18 WF-254S precoated silica TLC plates (0.25 mm) which were visualized using UV, ninhydrin, *p*-anisaldehyde, cerium ammonium molybdate, 2,4-dinitrophenylhydrazine, or bromocresol green stain. Flash column chromatography was performed using Machery-Nagel 60 μm (230-400 mesh) silica gel. All compounds purified by flash column chromatography were sufficiently pure for use in further experiments unless otherwise indicated. ^{1}H and ^{13}C NMR spectra were obtained in $CDCl_3$ or CD_3OD at rt (25 °C), unless otherwise noted, on Varian 400 MHz or Varian 600 MHz spectrometers. Chemical shifts of ^{1}H NMR spectra were recorded in parts per million (ppm) on the δ scale referenced to residual solvent peaks. Electrospray ionization liquid chromatography-mass spectrometry (LC-MS) analysis was performed on an Agilent G6545A quadrupole-time of flight mass spectrometer in positive mode with an Agilent 1290 UPLC system. Solvent A = water with 0.1% formic acid. Solvent B = 95% acetonitrile, 5% water and 0.1% formic acid. MS data were analyzed using the Agilent Qualitative Mass Hunter software.

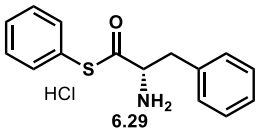
B. Compound synthesis



6.S1



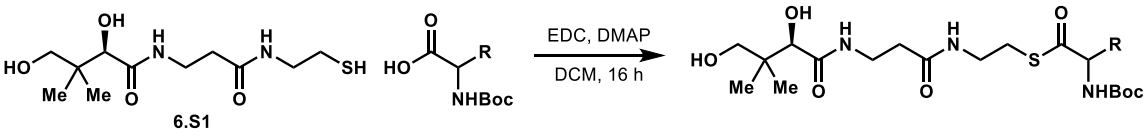
6.S2



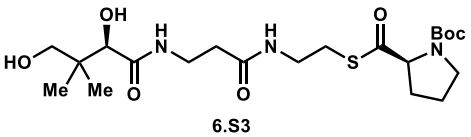
6.29

D-pantetheine (**6.S1**) was prepared according to the procedure described in Chapter 3. HSNAC (**6.S2**), and Phe-SPh (**6.29**),[38] and were kindly provided by the lab of Prof. David Sherman.

General procedure for the coupling of Boc-protected amino acids to D-pantetheine or HSNAC[39]



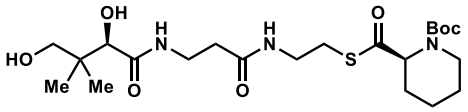
A 1 M solution of **6.S1** (187 mg, 0.67 mmol, 1.0 equiv) or HSNAC (72 µL, 0.67 mmol, 1.0 equiv) in dry DCM (6 mL) in a round bottomed flask was cooled to 0 °C in an ice bath. In a separate pear-shaped flask, Boc-protected amino acid (0.71 mmol, 1.05 equiv), EDC·HCl (129 mg, 0.67 mmol, 1.0 equiv) and DMAP (10 mg, 0.08 mmol, 0.13 equiv) were dissolved in 6 mL of DCM, and then transferred to the round bottomed flask via cannula. The ice bath was removed, and the reaction stirred overnight at rt under N_2. Solvent was removed under reduced pressure and the crude residue was purified by flash chromatography.



6.S3

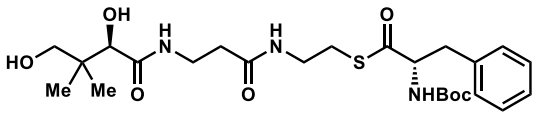
***tert*-butyl (*S*)-2-(((2-(3-((*R*)-2,4-dihydroxy-3,3-dimethylbutanamido)propanamido)ethyl)-thio) carbonyl)pyrrolidine-1-carboxylate [Boc-Pro-Pant] (6.S3).** The above procedure afforded 198 mg (86%) of the title compound as a clear syrup. Purified by flash chromatography (2-10%

gradient of MeOH in DCM). R_f = 0.3 (10% MeOH in DCM). **^{1}H NMR** (600 MHz, CD_3OD) δ 4.43 (dt, *J* = 9.4, 5.2 Hz, 1H), 3.93 (s, 1H), 3.56 – 3.47 (m, 6H), 3.36 (q, *J* = 6.7 Hz, 2H), 3.05 (t, *J* = 6.6 Hz, 2H), 2.48 – 2.43 (m, 2H), 2.39 – 2.23 (m, 1H), 2.07 – 1.91 (m, 3H), 1.52 (s, 4H), 1.45 (s, 5H), 0.96 (s, 9H); **^{1}H NMR** (400 MHz, $CDCl_3$) (approximately 2:1 ratio of rotamers) δ 7.70 – 7.53 (br s, 1H), 6.98 – 6.67 (m, 1H), 4.06 – 4.00 (m, 1H), 3.67 – 3.54 (m, 2H), 3.54 – 3.49 (m, 2H), 3.45 – 3.34 (m, 2H), 3.16-2.94 (m, 2H), 2.51-2.40 (m, 2H), 2.25 – 2.14 (m, 1H), 2.05 –1.84 (m, 3 H), 1.5 –1.39 (s, 9H), 1.09 – 0.90 (s, 6H); **^{13}C NMR** (150 MHz, CD_3OD) δ 203.1, 176.0, 173.7, 81.9, 70.3, 67.6, 47.7, 40.3, 36.3, 32.5, 31.7, 28.6, 24.3, 21.3, 20.9; **HRMS** (ESI) calcd for $C_{21}H_{38}N_3O_7S$ $[M+H]^+$ *m/z* 476.2425, found 476.2424.



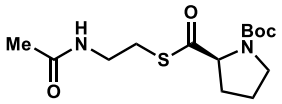
6.S4

***tert*-butyl (*S*)-2-(((2-(3-((*R*)-2,4-dihydroxy-3,3-dimethylbutanamido)propanamido)ethyl)-thio) carbonyl)piperidine-1-carboxylate (6.S4).** The above procedure afforded 189 mg (39%) of the title compound as a clear syrup. Purified by flash chromatography (2-10% gradient of MeOH in DCM). R_f = 0.25 (10% MeOH in DCM). **^{1}H NMR** (400 MHz, CD_3OD) δ 4.02 (dd, *J* = 16.0, 4.0 Hz,1H), 3.89 (s, 1H), 3.55 – 3.34 (m, 8H), 3.11 – 2.98 (m, 2H), 2.60 (t, *J* = 6.9 Hz, 1H), 2.42 (q, *J* = 6.7, 5.8 Hz, 2H), 2.37 – 2.25 (m, 1H), 1.65 (d, *J* = 11.6 Hz, 3H), 1.49 (d, 5H), 1.47 (s, 4H), 1.42 – 1.23 (m, 1H), 0.92 (s, 6H); **^{13}C NMR** (100 MHz, $CDCl_3$) δ 201.6, 174.0, 171.9, 110.2, 93.3, 81.1, 70.9, 50.9, 42.6, 39.6, 39.4, 36.0, 35.2, 28.5, 26.4, 24.7, 24.64, 21.8, 20.6; HRMS (ESI) calcd for $C_{22}H_{40}N_3O_7S$ $[M+H]^+$ *m/z* 490.2581, found 490.2587.



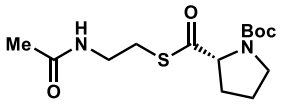
6.S5

***S*-(2-(3-((*R*)-2,4-dihydroxy-3,3-dimethylbutanamido)propanamido)ethyl) (*S*)-2-((*tert*-butoxycarbonyl)amino)-3-phenylpropanethioate (6.S5) [Boc-Phe-Pant]:** The above procedure afforded 141 mg (40%) of the title compound as a clear syrup. Purified by flash chromatography (2-10% gradient of MeOH in DCM). R_f = 0.3 (10% MeOH in DCM). **^{1}H NMR** (600 MHz, $CDCl_3$, mixture of rotamers) δ 7.64 – 7.44 (m, 2H), 7.31 – 7.05 (m, 5H), 6.99 (s, 1H), 5.41 – 5.14 (m, 1H), 4.57 – 4.42 (m, 1H), 3.98 (s, 1H), 3.51 (p, *J* = 6.3 Hz, 3H), 3.49 – 3.39 (m, 4H), 3.36 (q, *J* = 7.1 Hz, 2H), 3.04 – 2.90 (m, *J* = 9.9, 7.9 Hz, 2H), 2.61 (p, *J* = 7.1 Hz, 2H), 2.42 (dt, *J* = 30.2, 6.0 Hz, 2H), 1.42 – 1.22 (m, 9H), 0.94 (d, *J* = 7.5 Hz, 3H), 0.89 (s, 3H); **^{13}C NMR** (150 MHz, $CDCl_3$) δ 201.73, 174.30, 171.96, 155.61, 155.51, 129.48, 129.30, 128.76, 128.52, 127.21, 77.37, 77.16, 76.95, 70.65, 65.91, 61.57, 42.68, 42.65, 39.40, 39.36, 38.01, 35.81, 35.45, 28.36, 24.33, 21.46, 21.29, 20.64, 15.30.



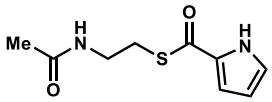
6.S7

***tert*-butyl (*S*)-2-(((2-acetamidoethyl)thio)carbonyl)pyrrolidine-1-carboxylate (6.S7) [Boc-Pro-SNAC]:** The above procedure afforded 610 mg (83% yield) of the title compound as an off-white powder. Purified by flash chromatography (5% MeOH in DCM). R_f = 0.5 (10% MeOH in DCM). **^{1}H NMR** (400 MHz, $CDCl_3$, mixture of rotamers) δ 6.09 (s, 1H), 4.43 (dd, *J* = 8.8, 3.1 Hz, 1H), 3.61 – 3.31 (m, 4H), 3.11 – 2.95 (m, 2H), 2.29 – 2.13 (m, 1H), 2.04 – 1.83 (m, 6H), 1.47 (s, 9H).



6.S8

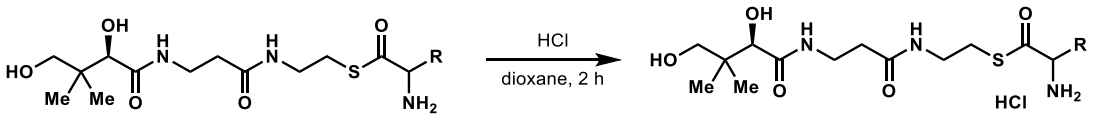
***tert*-butyl (*R*)-2-(((2-acetamidoethyl)thio)carbonyl)pyrrolidine-1-carboxylate (6.S8) [Boc-D--Pro-SNAC]:** The above procedure afforded 14 mg (5% yield) of the title compound as an off-white powder. Purified by flash chromatography (5% MeOH in DCM). R_f = 0.5 (10% MeOH in DCM). **^{1}H NMR** (400 MHz, $CDCl_3$, mixture of rotamers) δ δ 5.95 (s, 1H), 4.39 (ddd, *J* = 29.1, 8.8, 3.3 Hz, 1H), 3.41 (overlapping m and q, *J* = 6.2 Hz, 4H), 3.01 (t, *J* = 6.4 Hz, 2H), 2.34 (s, 3H), 1.96 (s, 4H), 1.49 – 1.11 (m, 9H).



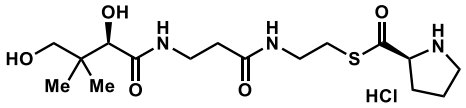
6.S9

***S*-(2-acetamidoethyl) 1H-pyrrole-2-carbothioate (6.S9):** Used the unprotected acid. The above procedure afforded 582mg (57% yield) as a light-green solid. Purified by flash chromatography (2% MeOH in DCM). R_f = 0.3 (5% MeOH in DCM). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 9.29 (s, 1H), 7.04 (tt, *J* = 3.7, 1.7 Hz, 2H), 6.29 (dt, *J* = 3.7, 2.5 Hz, 1H), 6.05 (s, 1H), 3.52 (q, *J* = 6.0 Hz, 2H), 3.19 (dd, *J* = 6.8, 5.7 Hz, 2H), 1.97 (s, 3H). All spectra obtained matched literature values.[40]

General procedure for the deprotection of Boc-protected thioesters

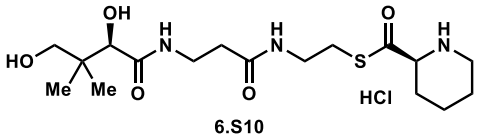


To a 1-dram vial containing Boc-protected aminoacyl thioesters was added excess HCl (4 M in dioxane). The reaction mixture stirred sealed for 2 h, before concentrating and washing 3x with hexanes. Products were isolated as hydrochloride salts and stored as 100 mM solutions in DMSO at -20 °C until further use.



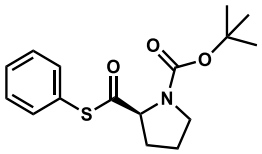
6.21

***S*-(2-(3-((*R*)-2,4-dihydroxy-3,3-dimethylbutanamido) propanamido)ethyl) (*S*)-pyrrolidine-2-carbothioate, hydrochloride salt [Pro-Pant] (6.21).** The above procedure afforded 38 mg (100%) of the title compound as a white paste. R_f = 0 (10% MeOH in DCM). **^{1}H NMR** (600 MHz, CD_3OD) δ 4.66 (q, *J* = 5.9, 4.1 Hz, 1H), 4.15 (s, 1H), 3.49 – 3.37 (m, 6H), 2.63 (t, *J* = 6.5 Hz, 2H), 2.54 – 2.41 (m, 2H), 2.17 – 2.03 (m, 4H), 1.16 (s, 3H), 1.00 (s, 3H); **^{13}C NMR** (150 MHz, CD_3OD) δ 196.8, 179.2, 172.4, 7.15, 76.5, 67.2, 49.8, 49.4, 49.2, 49.1, 49.0, 48.8, 48.7, 48.5, 47.1, 41.7, 39.5, 37.1, 32.8, 30.7, 29.7, 24.5, 22.7, 19.2; HRMS (ESI) calcd for $C_{16}H_{30}N_3O_5S$ $[M+H]^+$ *m/z* 376.1901, found 376.1904.



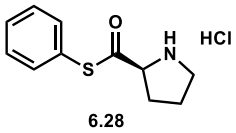
6.S10

***S*-(2-(3-((*R*)-2,4-dihydroxy-3,3-dimethylbutanamido) propanamido)ethyl) (*S*)-piperidine-2-carbothioate, hydrochloride salt (6.S10):** The above procedure afforded 49 mg (100%) of the title compound as a white paste. R_f = 0 (10% MeOH in DCM). **^{1}H NMR** (600 MHz, CD_3OD) δ 4.29 (dd, *J* = 11.2, 3.7 Hz, 1H), 4.16 (s, 1H), 3.43 (t, *J* = 6.2 Hz, 8H), 2.64 (t, *J* = 6.5 Hz, 2H), 2.47 – 2.30 (m, 2H), 1.98 – 1.69 (m, 6H), 1.18 (s, 3H), 1.01 (s, 3H); **^{13}C NMR** (150 MHz, CD_3OD) δ 197.6, 179.2, 172.4, 77.1, 76.5, 64.5, 49.8, 45.3, 41.7, 39.6, 32.8, 29.3, 28.8, 22.8, 19.2; HRMS (ESI) calcd for $C_{17}H_{32}N_3O_5S$ $[M+H]^+$ *m/z* 390.2057, found 390.2063.



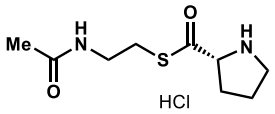
6.S11

***Tert*-butyl (*S*)-2-((phenylthio)carbonyl)pyrrolidine-1-carboxylate [Boc-Pro-SPh] (6.S11)**.[41] A solution of Boc-amino acid (1.0 equiv) and diphenyl disulfide (1.5 equiv) in DCM (40x) was cooled to 0 °C in an ice bath. Tributylphosphine (1.5 equiv) was added dropwise over 5 min. The reaction mixture stirred under N_2 for 1 h until completion, as visualized by TLC. The reaction was quenched with sat. aq. $NaHCO_3$ and extracted 3x with DCM. The combined organic layers were dried over Na_2SO_4, filtered and concentrated. The crude residue was purified by flash chromatography. R_f = 0.25 (10% EtOAc in hexanes). The crude mixture was purified by column chromatography using in 10% EtOAc in hexanes. Isolated 0.59 g (82%) as a white crystalline solid. **^{1}H NMR** (600 MHz, $CDCl_3$) (approximately 2:1 mixture of rotamers) δ 7.48 – 7.35 (m, 5H), 4.59 – 4.43 (dd, *J* = 8.8, 3.6 Hz, 1H), 3.68 – 3.41 (m, 2H), 2.34 – 1.90 (m, 4H), 1.53 – 1.47 (s, 9H); ^{13}C NMR (150 MHz, $CDCl_3$) δ 200.8, 154.0, 134.6, 129.4, 129.3, 127.7, 80.8, 66.3, 46.8, 31.7, 28.6, 23.7. All spectra obtained were consistent with literature values.[42]



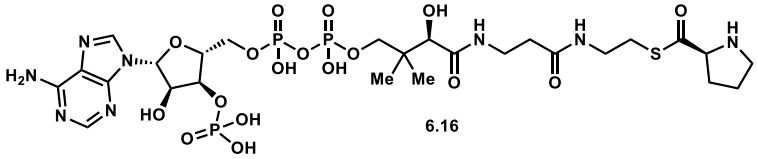
6.28

***S*-phenyl (*S*)-pyrrolidine-2-carbothioate hydrochloride salt [Pro-SPh(·HCl)] (6.28)**: The deprotection procedure afforded 79 mg (100%) of the title compound from 100 mg of substrate as a white paste. **^{1}H NMR** (CD_3OD, 600 MHz) δ 7.70 – 7.32 (m, 5H), 4.82 – 4.78 (m, 1H), 3.39 (ddt, *J* = 11.6, 7.3, 3.6 Hz, 2H), 2.56 (dddd, *J* = 13.3, 8.5, 7.6, 5.9 Hz, 1H), 2.22 (dq, *J* = 13.1, 7.6 Hz, 1H), 2.17 – 2.04 (m, 2H); ^{13}C NMR (CD_3OD, 150 MHz) δ 195.4, 136.0, 131.5, 130.8, 126.45, 67.1, 47.31, 30.8, 24.8; HRMS (ESI) calcd for $C_{11}H_{13}NOS$ $[M+H]^+$ *m/z* 208.0791, found 208.0790.



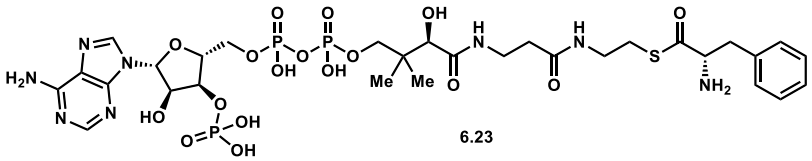
6.S12

S-(2-acetamidoethyl) (*R*)-pyrrolidine-2-carbothioate (6.S12) [D-Pro-SNAC]: The deprotection procedure afforded 11 mg (quantitative yield) of the title compound from 14 mg of substrate as an off-white powder. **^{1}H NMR** (400 MHz, CD_3OD) δ 4.64 (t, *J* = 7.7 Hz, 1H), 3.78 – 3.55 (m, 1H), 3.42 (td, *J* = 6.6, 3.8 Hz, 1H), 3.34 (t, *J* = 6.7 Hz, 2H), 3.26 – 3.10 (m, 1H), 3.01 (t, *J* = 6.7 Hz, 2H), 2.55 – 2.45 (m, 1H), 2.33 (s, 3H), 2.21 – 2.03 (m, 2H).



6.16

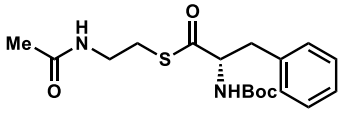
Prolyl-coenzyme A [Pro-CoA] (6.16): Aminoacyl thioester (1.9 mg, 0.0076 mmol, 3.0 equiv) and coenzyme A, trilithium salt (2.0 mg, 0.0025 mmol, 1.0 equiv) were dissolved in 2.0 mL 50 mM phosphate buffer, pH 8.5 in a 1 dram vial. The reaction mixture stirred sealed for 30 min and was filtered through a 0.2 µm syringe filter and washed with 0.2 mL of MeOH for purification by preparative HPLC in one injection (Beckman Coulter 126 Solvent Module/166 detector at 260 nm; Phenomenex Luna C18(2) column, 250 x 22 mm, 5 µm particle size; the flow rate was 8 mL/min, solvent A was 100 mM ammonium formate lowered to pH 5 with formic acid and solvent B was acetonitrile). Product was eluted using a method of 10 min at 5% B, linear gradient to 95% B over 20 min, 9 min at 95% B, linear gradient to 5% B over 1 min, hold for 15 min at 5% B. Relevant fractions were pooled and concentrated to remove acetonitrile and then lyophilized. Purified aminoacyl-CoA compounds were stored as 1 mM solutions in water at -20 °C until further use. t_R = 21 min. Isolated 3 mg (quantitative) as a white powder. **HRMS** (ESI) calcd for $C_{26}H_{44}N_8O_{17}P_3S$ $[M+H]^+$ *m/z* 865.1752, found 865.1724.



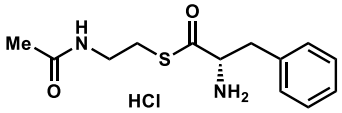
6.23

Phe-CoA (6.23): The above procedure afforded 1.2 mg (52%) of the title compound as a white powder. t_R = 24 min. **HRMS** (ESI) calcd for $C_{30}H_{46}N_8O_{17}P_3S$ $[M+H]^+$ *m/z* 915.1909, found 915.1907.

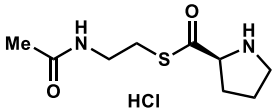
The following compounds were synthesized by Joshua Pyser:



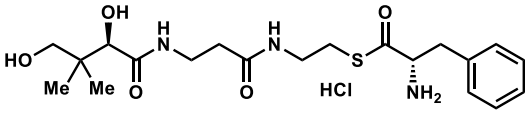
6.S13



6.32

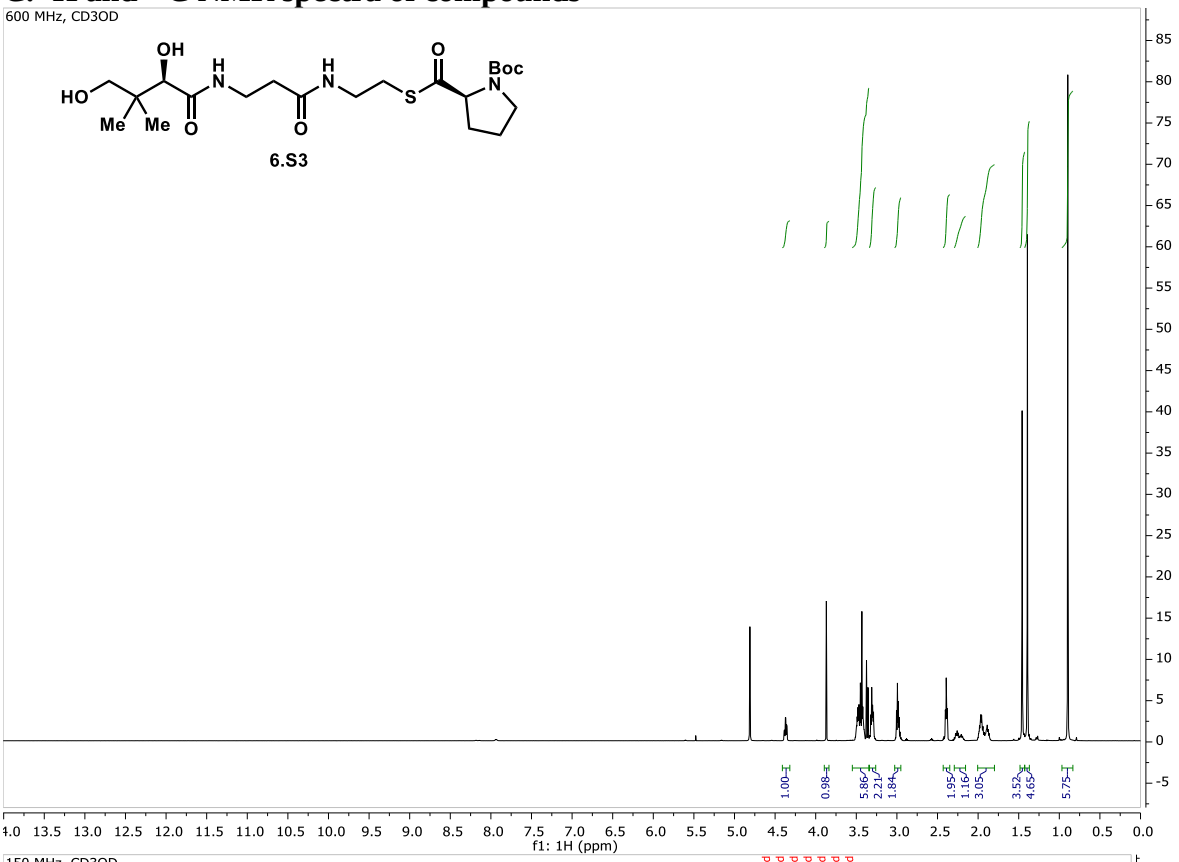


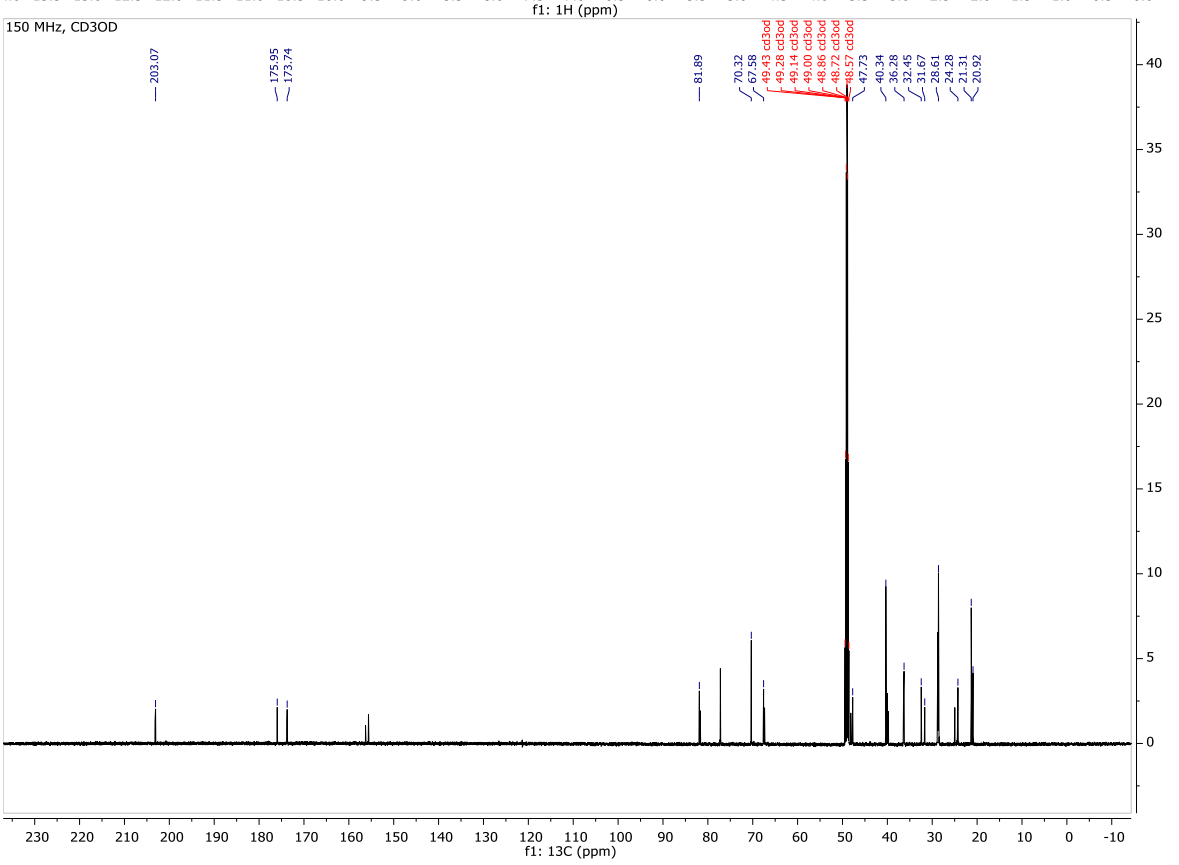
6.31

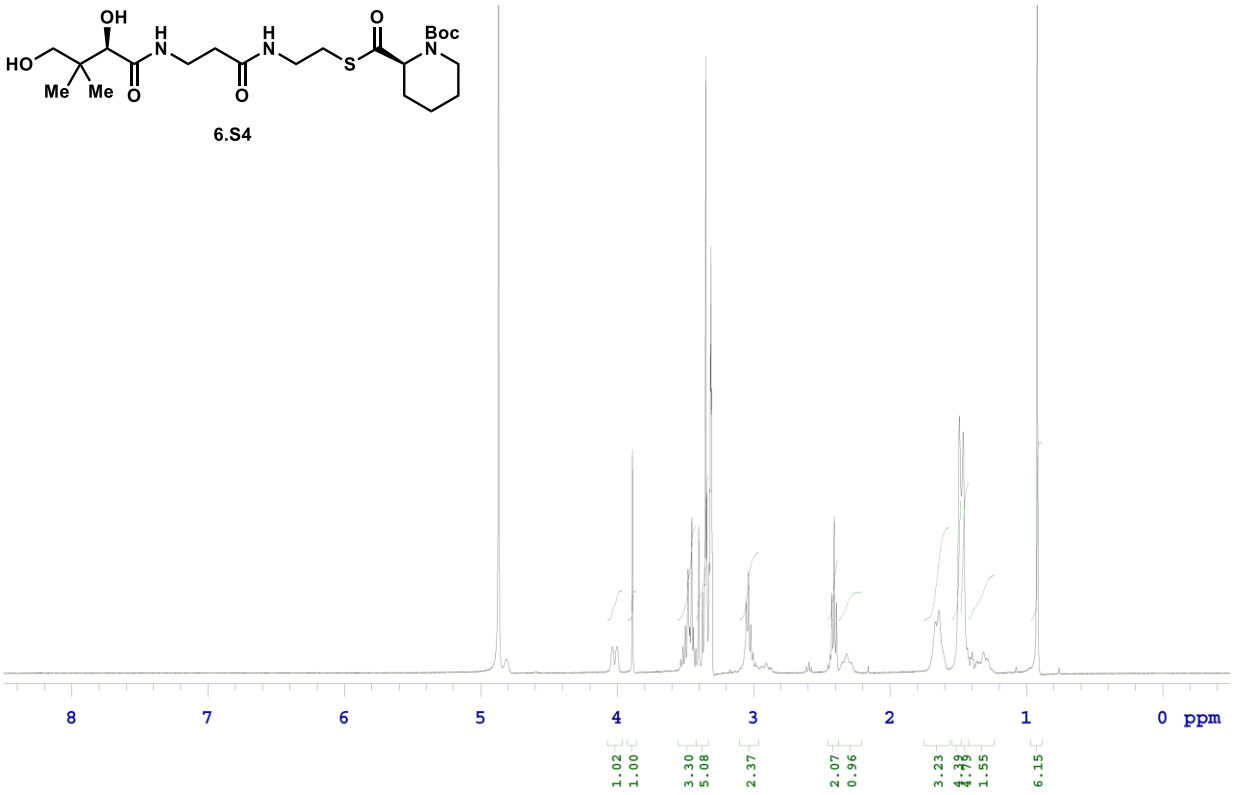


6.22

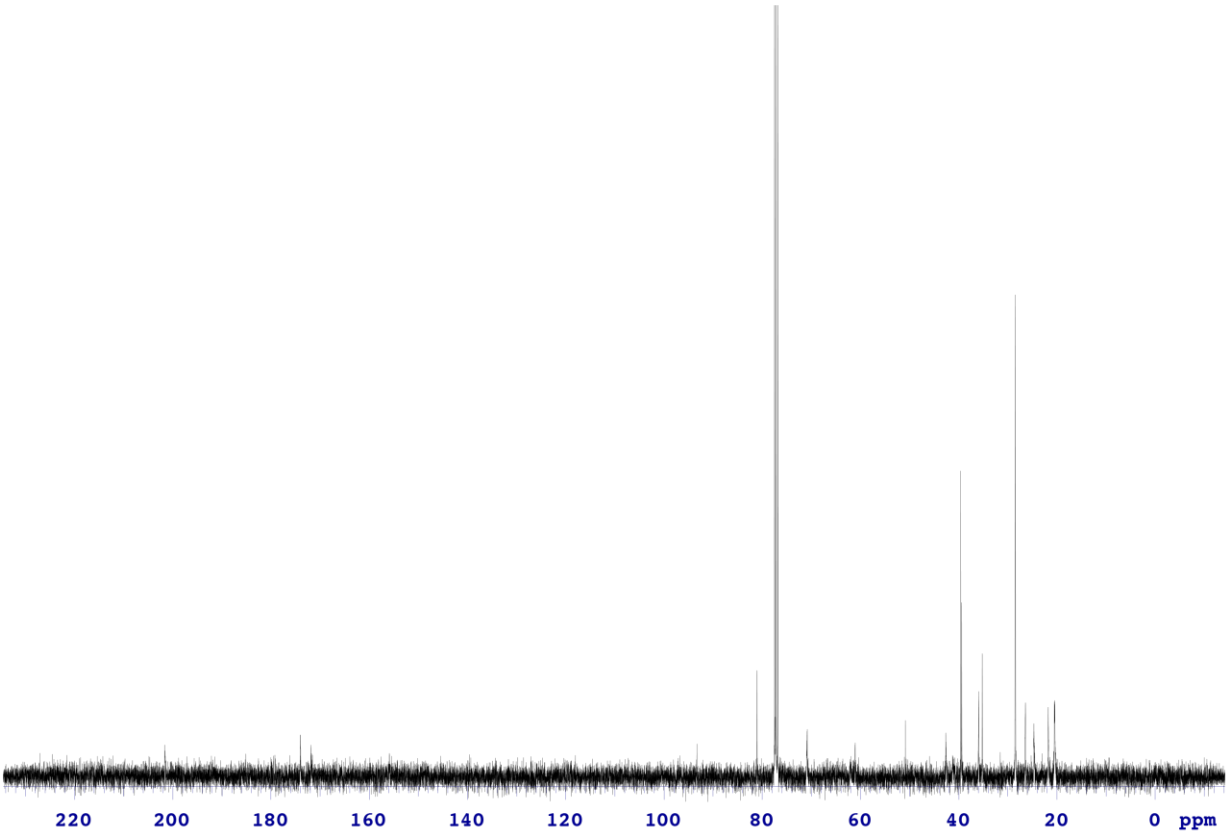
C. ^{1}H and ^{13}C NMR spectra of compounds

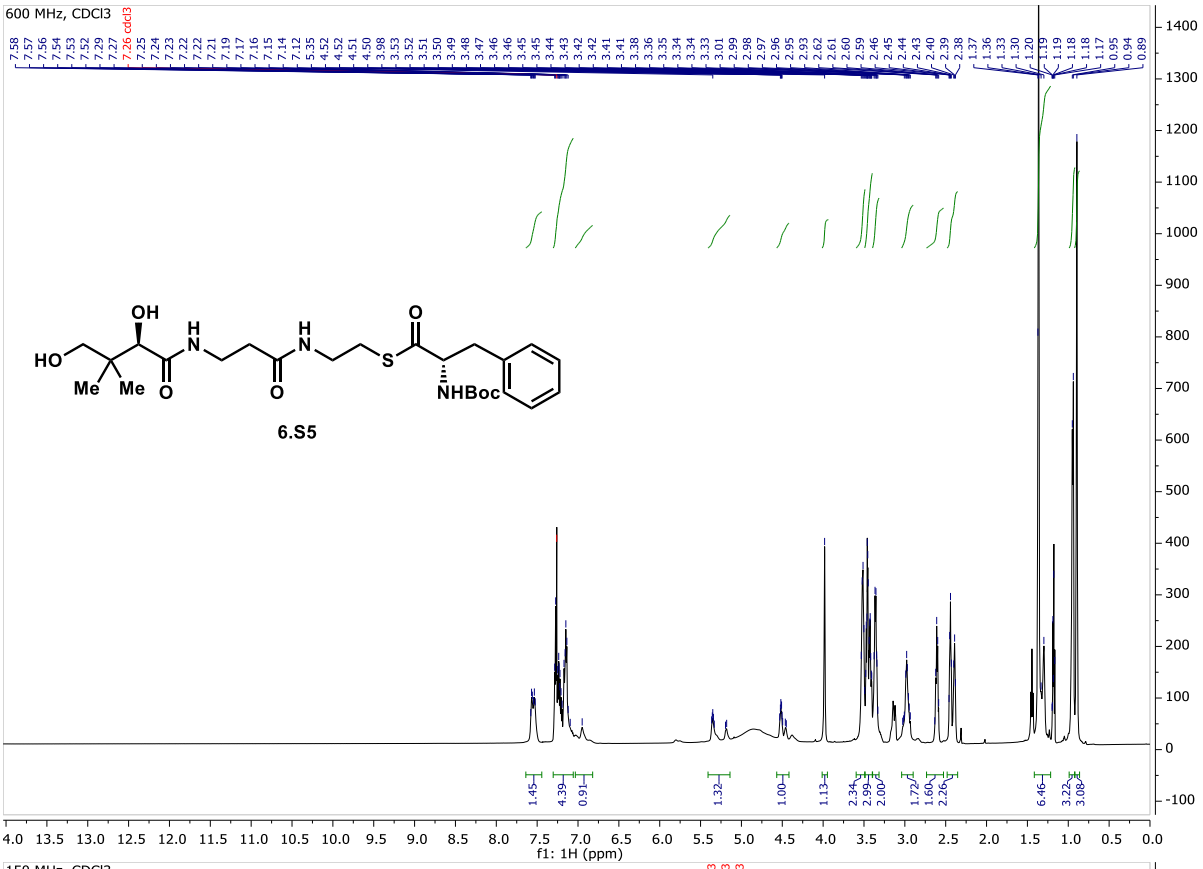




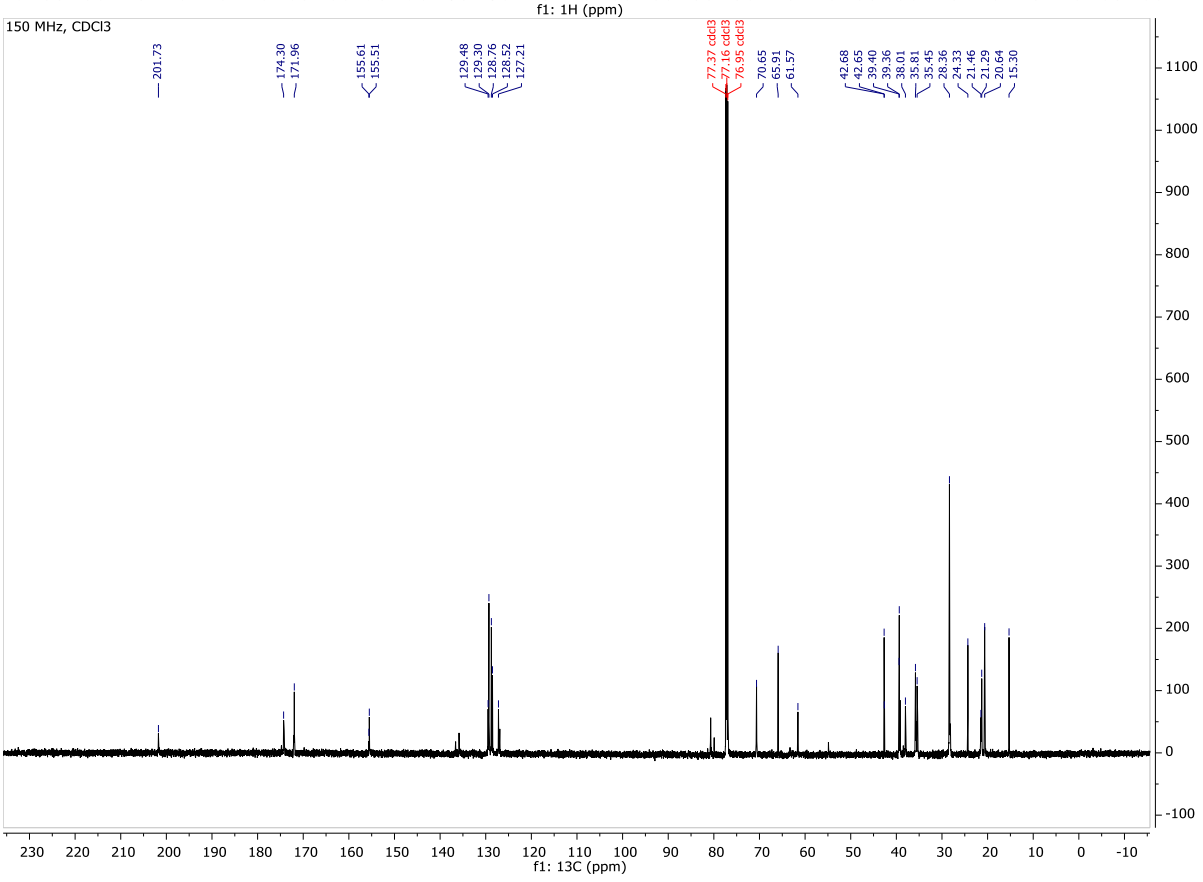

OH
HO
Me
Me
O
H
N
O
H
N
O
S
O
Boc
N
8
7
6
5
4
3
2
1
0 ppm
1.02
1.00
3.30
5.08
2.37
2.07
0.96
3.23
1.55
6.15

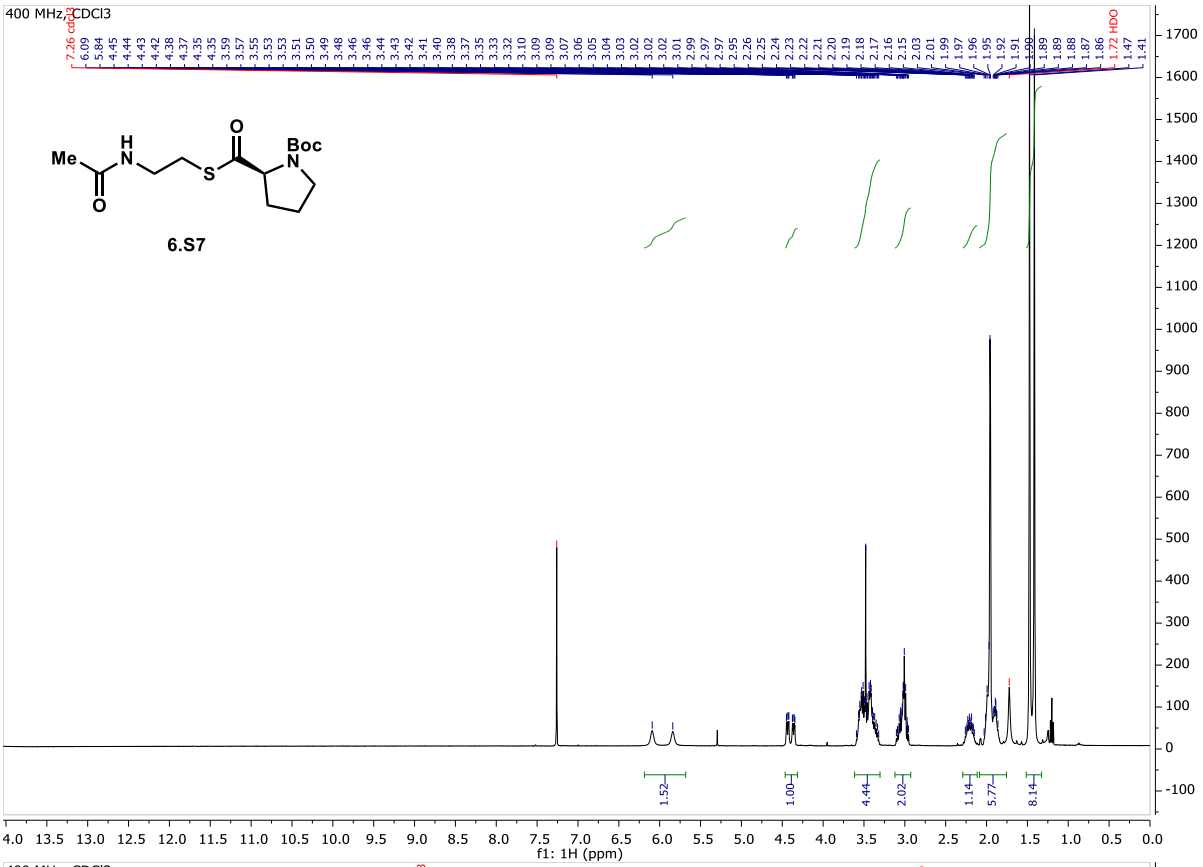
6.S4

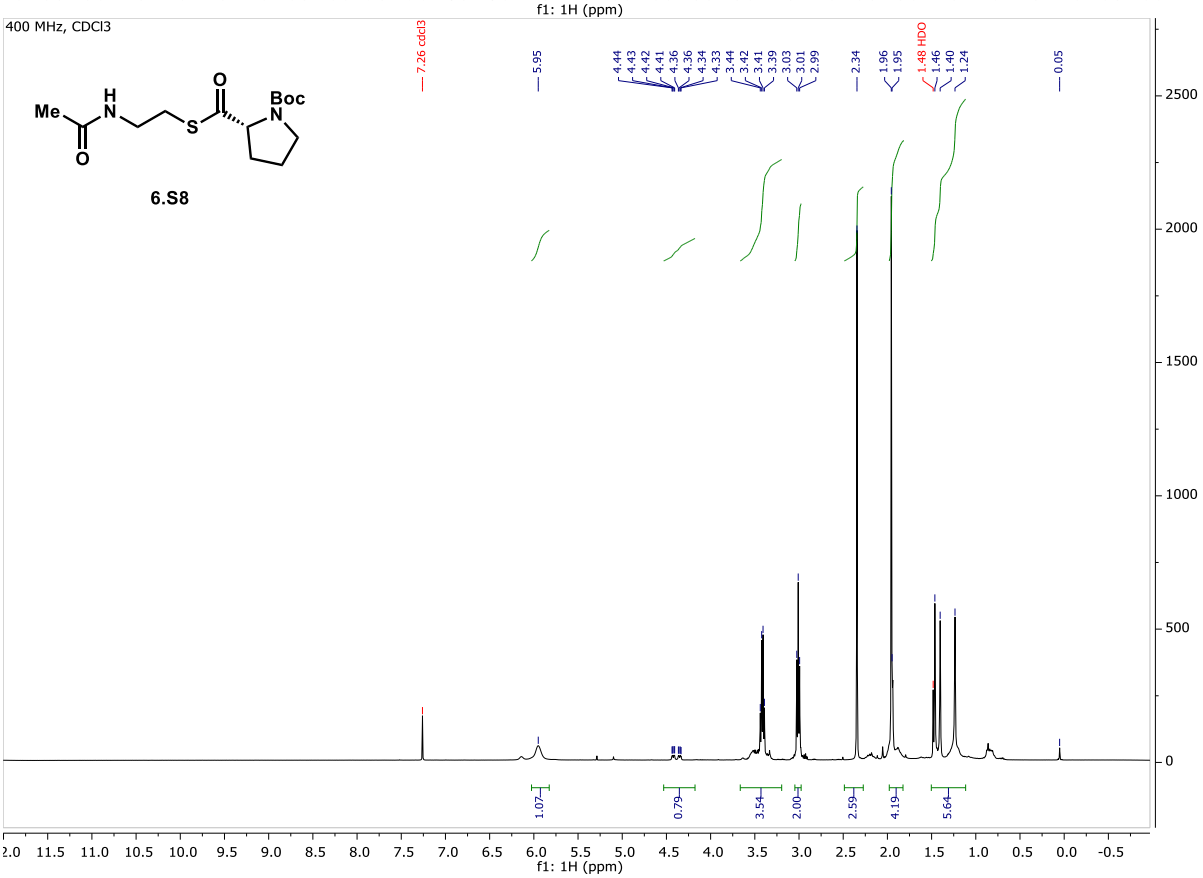

220
200
180
160
140
120
100
80
60
40
20
0 ppm

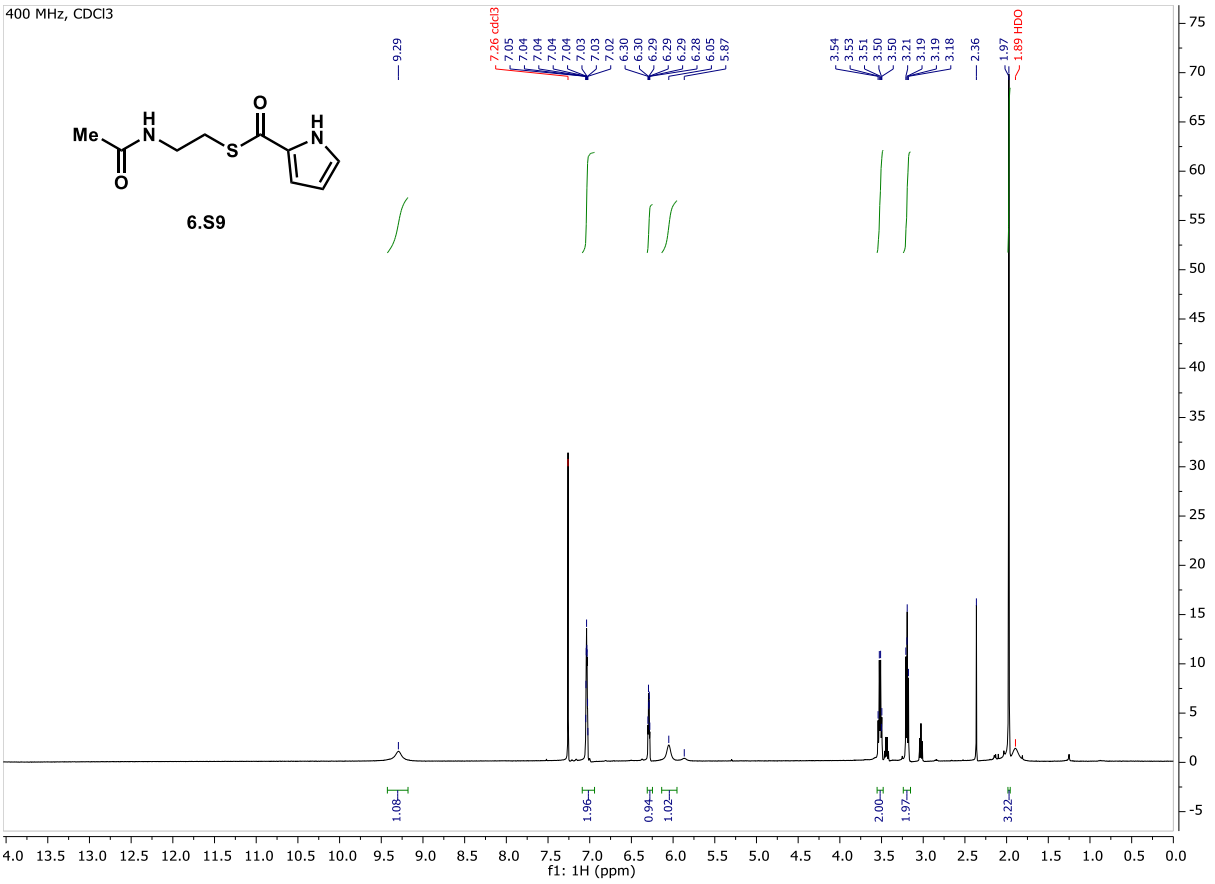


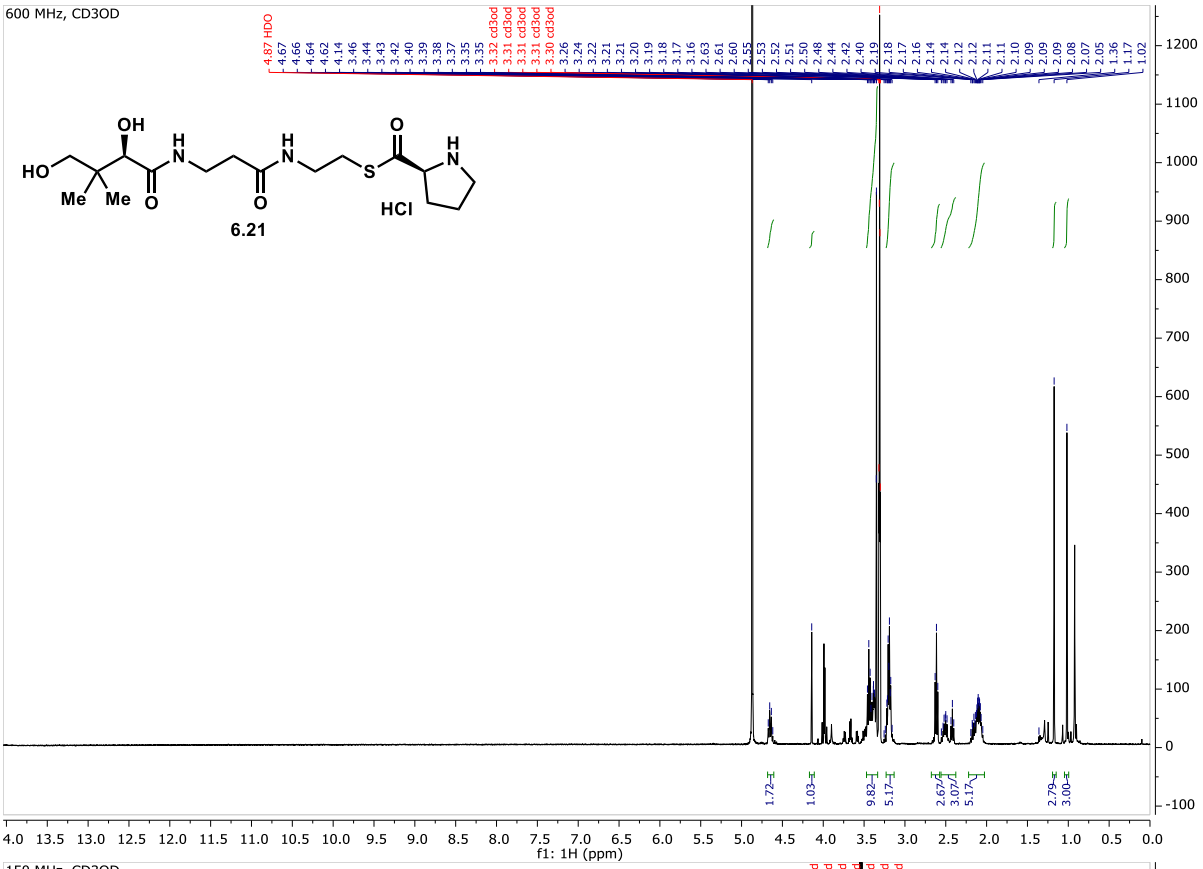
6.S5

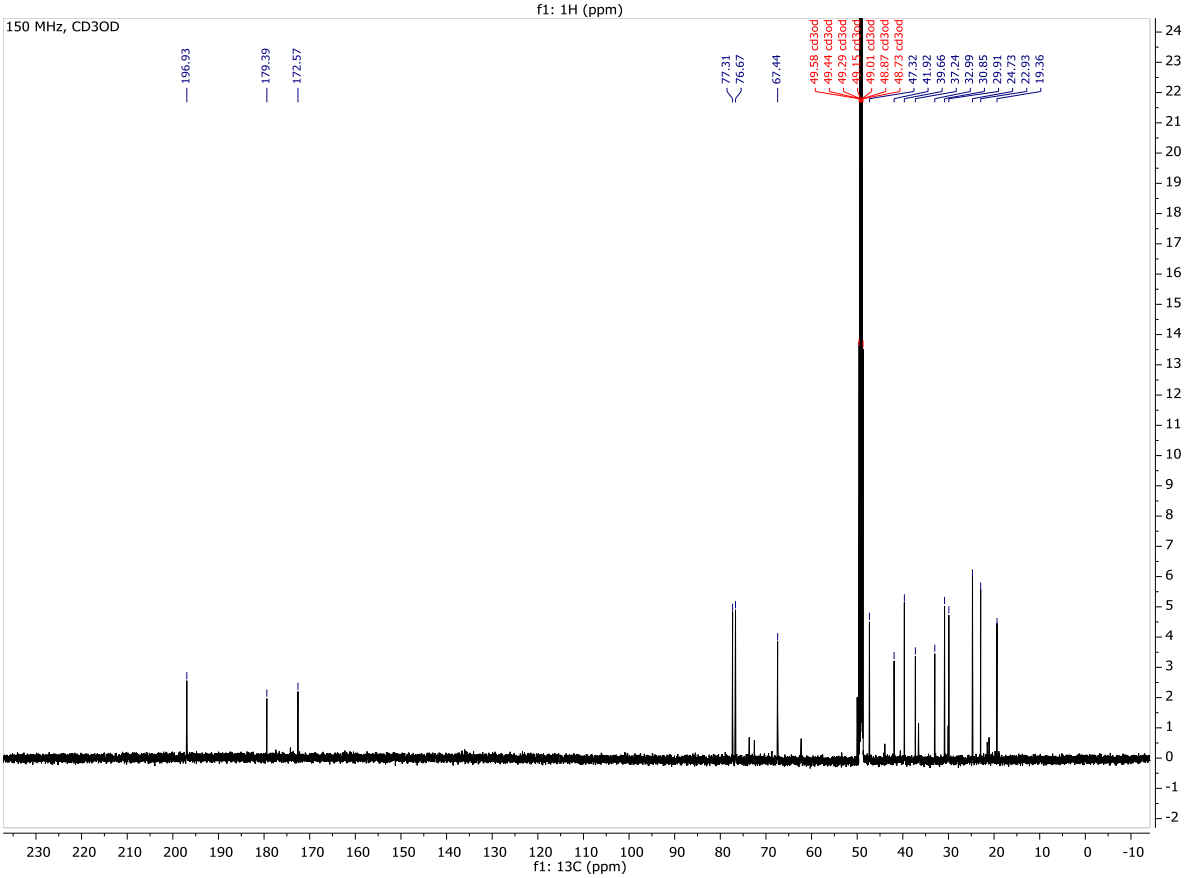


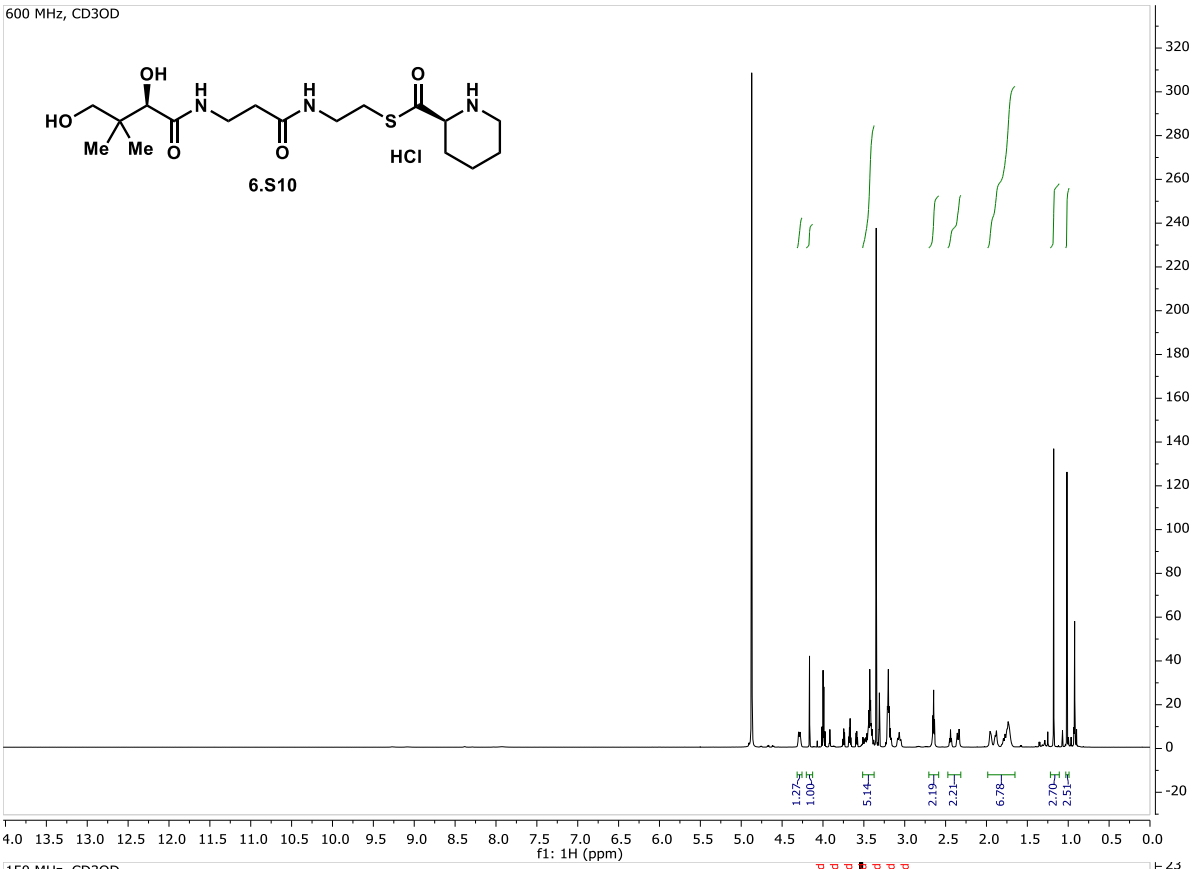

400 MHz, CDCl3
7.26 cdcl3
1.72 HDO
Me
H
N
O
S
O
Boc
N
6.S7
f1: 1H (ppm)

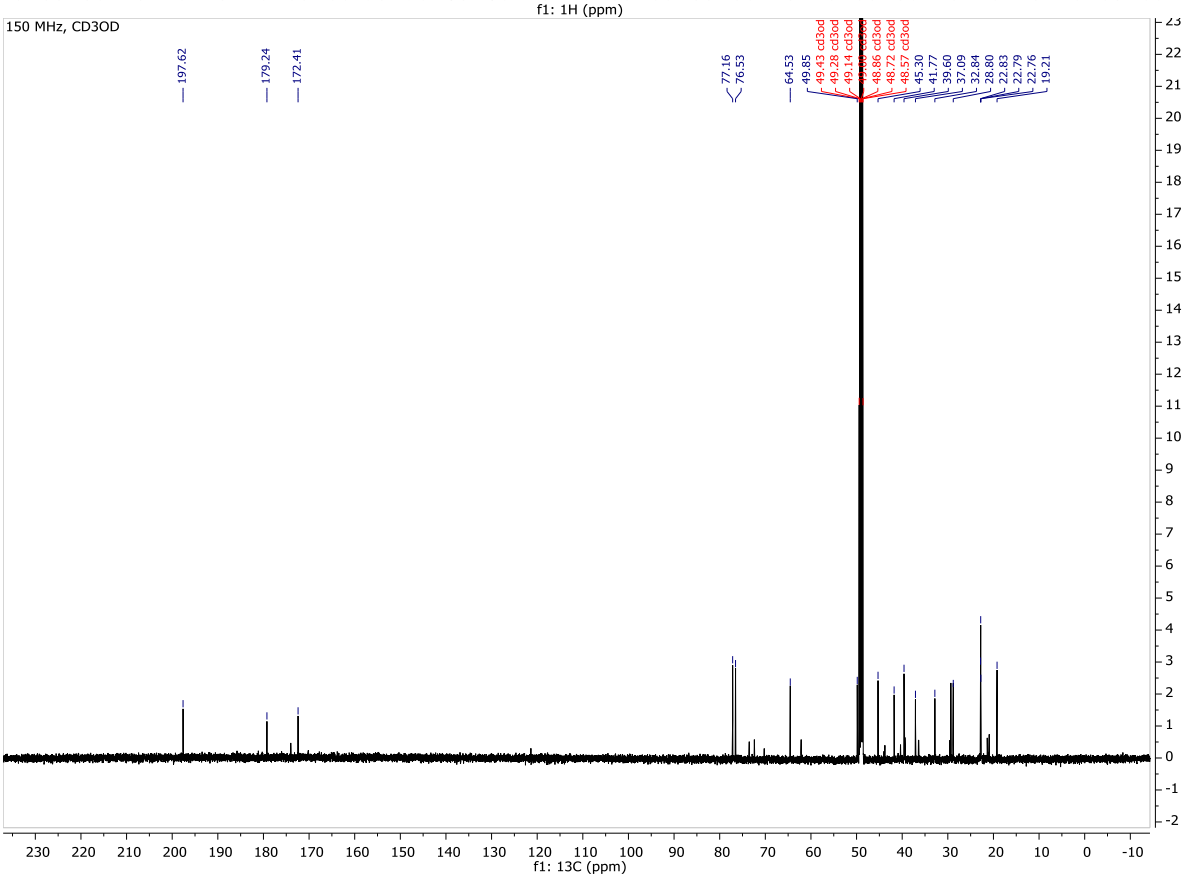

400 MHz, CDCl3
7.26 cdcl3
5.95
1.48 HDO
2.34
0.05
Me
H
N
O
S
O
Boc
N
6.S8
f1: 1H (ppm)

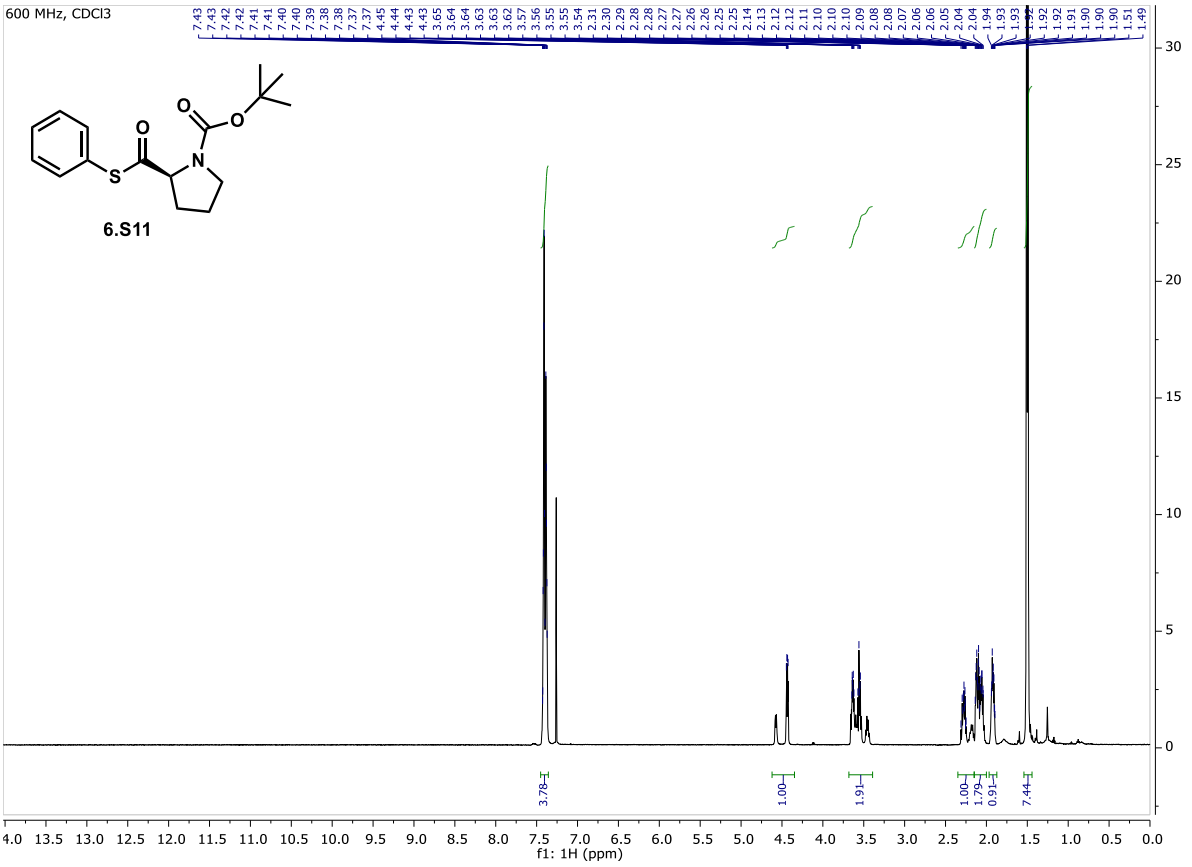

400 MHz, CDCl3
6.S9
f1: 1H (ppm)

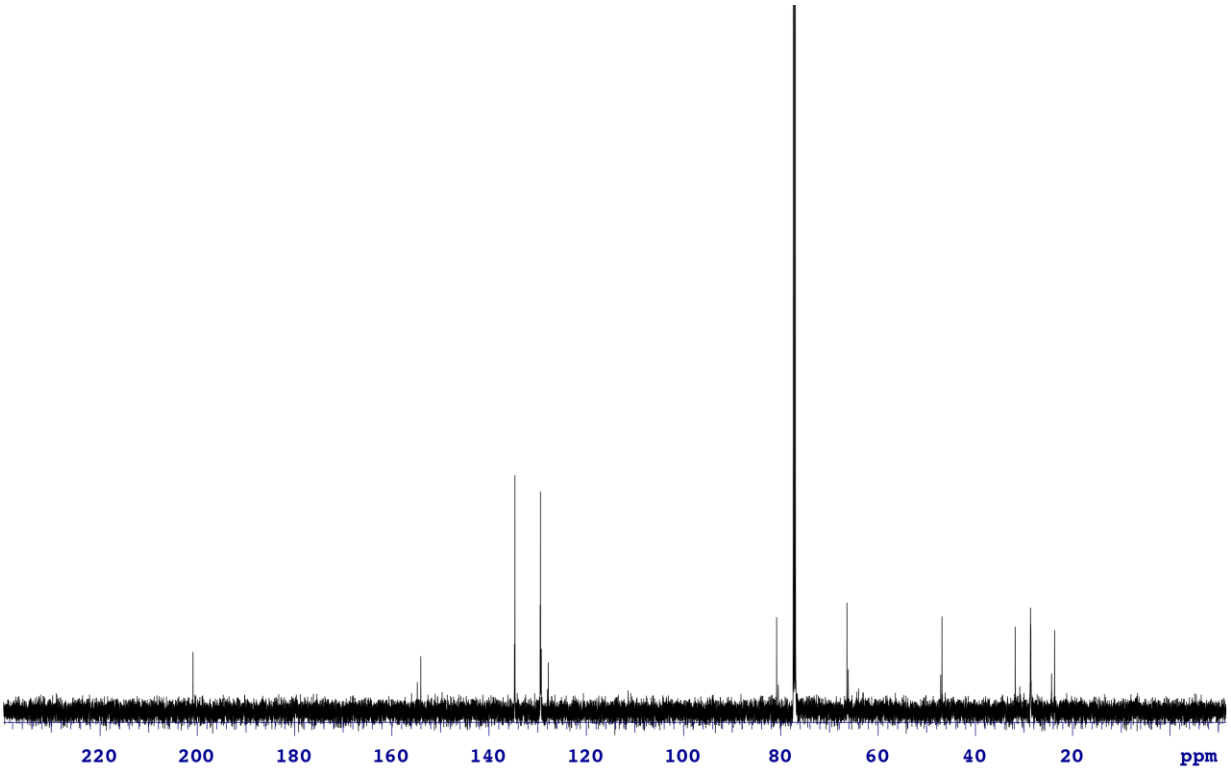

600 MHz, CD3OD
6.21
HCl
f1: 1H (ppm)

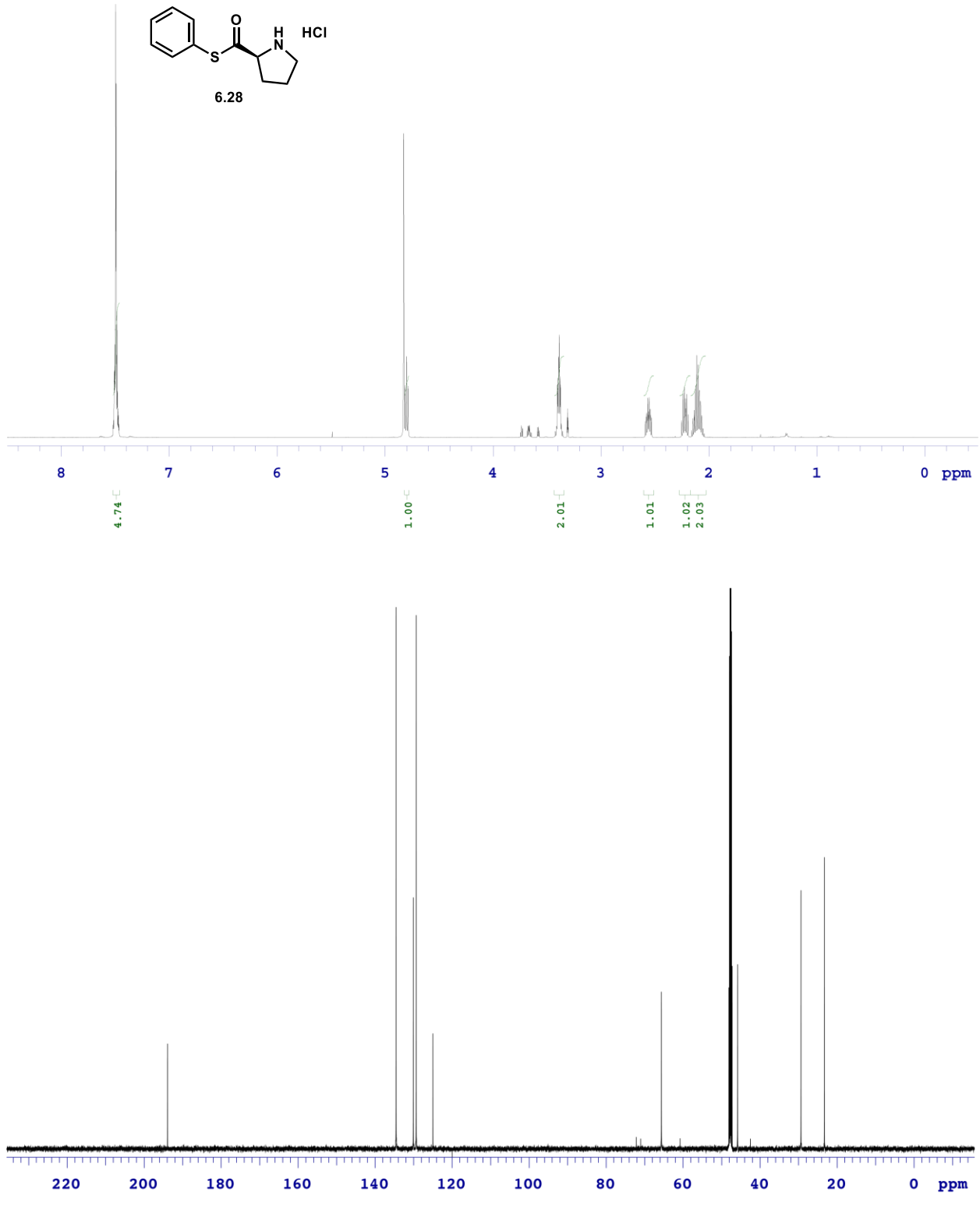

150 MHz, CD3OD
f1: 13C (ppm)

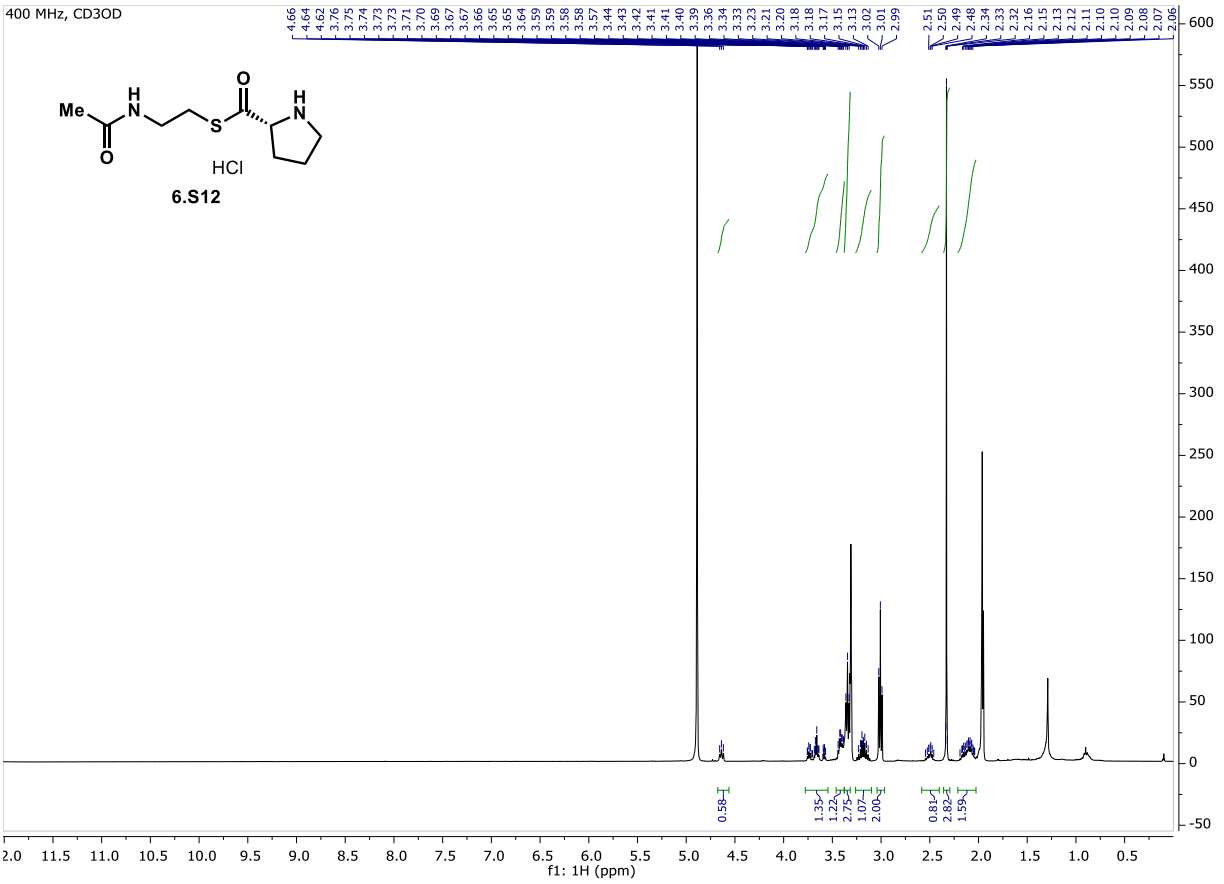

600 MHz, CD3OD
6.S10
f1: 1H (ppm)


150 MHz, CD3OD
197.62
179.24
172.41
77.16
76.53
64.53
49.85
f1: 13C (ppm)


600 MHz, CDCl3
6.S11
f1: 1H (ppm)


220
200
180
160
140
120
100
80
60
40
20
ppm


O
H
N
HCl
S
6.28
8
7
6
5
4
3
2
1
0 ppm
4.74
1.00
2.01
1.01
1.02
2.03
220
200
180
160
140
120
100
80
60
40
20
0 ppm


400 MHz, CD3OD
Me
H
N
O
S
O
H
N
HCl
6.S12
f1: 1H (ppm)

II. Cloning, protein expression and purification

A. General information

Escherichia coli cloning strain DH5α (Invitrogen) was used for DNA propagation. Protein was expressed in *E. coli* strains BL21, BL21(DE3) (New England BioLabs) or BAP1 after transformation with relevant plasmids. Cells were grown in Luria Bertani (LB) or Terrific broth (TB) with 4% v/v glycerol supplemented with kanamycin (50 μg/mL) purchased from Gold Biotechnology.

Primers were purchased from Integrated DNA Technologies and subcloned into various ligation-independent cloning (LIC) vectors. DpnI restriction enzyme was purchased from New England BioLabs. LIC-qualified T4 DNA polymerase was purchased from EMD Millipore. QIAquick PCR purification, gel extraction and miniprep kits were purchased from Qiagen. HisPur nickel-nitrilotriacetic acid (Ni-NTA) resin was purchased from Thermo Scientific. Proteins were concentrated using Amicon centrifugal filters purchased from EMD Millipore at 4,000 x *g*, 4 °C. PD-10 desalting columns were purchased from GE Healthcare. Protein samples were analyzed on Mini-PROTEAN TGX Gels (4-15%) from BioRad and visualized with Protein Ark Quick Coomassie Stain from Anatrace. Proteins were quantified with the Pierce 660 nm Assay Reagent from Thermo Scientific.

B. Sequence information

A pET28a plasmids containing *anaB* and *anaD* were obtained from the lab of Prof. Olivier Ploux at the French National Centre for Scientific Research, and were purified according to the published procedures.[9,10] Plasmids containing maltose binding protein fusions of *coaA*, *coaD* and *coaE* were

obtained from the lab of Prof. Michael Burkart of the University of California in San Diego. pET22b-*PaAcpH* was also obtained from the lab of Michael Burkart and was expressed and purified according to a published procedure.[18] Plasmids containing *rifR* and *redJ* were obtained from the laboratory of Prof. Janet Smith at the University of Michigan, and were expressed and purified according to published procedures.[27,28] A pET24b vector encoding the gene for the promiscuous phosphopantetheinyl-transferase Sfp (*Bacillus subtilis*)[17] was kindly donated by the group of Prof. David Sherman (University of Michigan Life Sciences Institute) and was expressed and purified according to the procedure in Chapter 2.

anaA from *Oscillatoria* PCC 6506 and *tycF* from *Brevibacillus brevis* subcloned into pET151 plasmids were purchased from Integrated DNA Technologies. Gene fragment of *P. fluorescens Pf_AcpH* and thioestersases GrsT, PltG and SrfAD (not used yet) were purchased from IDT.

Table 6.S1. Proteins used in this Chapter

Protein name	Source organism	GenBank accession
AnaD	*Oscillatoria* PCC 6506	CBN59194.1
AnaB	*Oscillatoria* PCC 6506	CBN59192.1
AcpP	*E. coli*	AAA23740.1
Pa_AcpH	*Pseudomonas aeruginosa*	CEI16732.1
Pf_AcpH	*Pseudomonas fluorescens*	KIR20923.1
AnaA	*Oscillatoria* PCC 6506	CBN59190.1
RedJ	*Streptomyces coelicolor* ATCC A3(2)	AL939125.1
RifR	*Amycolatopsis mediterranei*	AF040570.3
TycF	*Brevibacillus brevis*	AP008955.1
GrsT	*Aneurinibacillus migulanus*	AAA58717.1
PltG	*Pseudomonas fluorescens* Pf-5	AAY92065.1
SrfAD	*Bacillus subtilis*	CAA49819.1
Sfp	*Bacillus subtilis*	BAA09125.1

Pf_AcpH E. coli-optimized gene fragment:

```
ATGAACTATCTGGCACATCTGCATCTGGGTGGTCAGCTGCCTGCACAGCTGCTGGGTAGCCTGTATGGTGATTTTGT
TAAAGGTCGTCTGCAGGGTCAGTTTAGTCCGCAGATTGAAGCAGCAATTCAGCTGCATCGTAGCATTGATCGTTTTA
CCGATAGCCATCCGCTGGTTGGTGAAGCACTGAGCCGTTTTAGCCAGACCCGTCGTCGTTATGCAGGTATTGTTCTG
GATGTGTTTTTTGATCATTGTCTGGCACGTGATTGGGCACTGTATGCAGATCAGCCGCTGGAACGTTTTACCAGCCA
TGTTTATCAGGTTCTGGCAGCAGAACCGGCACTGCCTGGTCGTCTGGCACAGATTGCACCGTATATGGCAGCCGATG
ATTGGCTGGGTAGTTATCGTGAATTTGCAGTTATGGAACAGGTTCTGCGTGGTATTAGCCGTCGTCTGACCCAGCCG
GAAGAACTGGGTTATGCAATGCAAGAACTGCGTGTTCTGTATGAACCGCTGAGCGAAGATTTTCGTCTGTTTTATCC
GGAACTGCAGGCATTCGCACTGCAGTTTTAA
```

Translation:

```
MNYLAHLHLGGQLPAQLLGSLYGDFVKGRLQGQFSPQIEAAIQLHRSIDRFTDSHPLVGEALSRFSQTRRRYAGIVL
DVFFDHCLARDWALYADQPLERFTSHVYQVLAAEPALPGRLAQIAPYMAADDWLGSYREFAVMEQVLRGISRRLTQP
EELGYAMQELRVLYEPLSEDFRLFYPELQAFALQF*
```

For cloning into pMCSG7, pMCSG9 and pMocr (LIC overhangs in **bold**):

PfAcpH LIC fwd1 5′-**tacttccaatccaatgca**AACTATCTGGCACATCTGCATCTGGG-3′

PfAcpH LIC rev1 5′-**ttatccacttccaatgtta**AAACTGCAGTGCGAATGCCTG-3′

The LIC insert was amplified in a 50 µL reaction containing 10 µL Xtreme buffer, 1 µM gene fragment, 3 µM each of forward and reverse LIC primers, 200 µM each of dNTPs, 0.04 U/µL KOD Xtreme Hot Start polymerase (Millipore). Excised domains were amplified according to the following PCR procedure: 94 °C 2:00, (98 °C 0:10, 68 °C 1:00) for 39 cycles, 68 °C 5:00. The insert was purified by gel extraction and subcloned into pMCSG7, pMCSG9 and pMocr using standard LIC protocols.[43]

Pf_AcpH was subcloned into pTRC9 by Dr. Gregory Dodge to produce an MBP-fused protein.

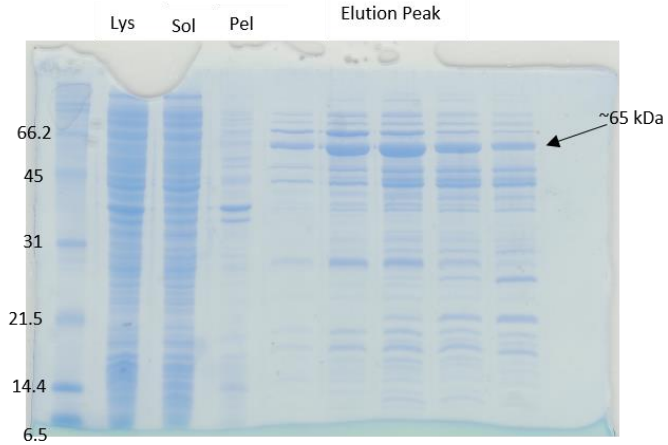


Figure 6.S1. SDS-PAGE gel of the Pf_AcpH-MBP fusion protein (expected MW: 65 kDa). Image courtesy of Dr. Gregory Dodge.

C. Protein expression and purification

Overexpression of CoaADE: ampicillin-resistant plasmids containing the gene of interest were transformed into chemically competent BL21(DE3) *E. coli* cells. A single colony was picked to inoculate a 5 mL LB starter culture supplemented with ampicillin (100 μg/mL) grown overnight at 37 °C, 200 rpm. The following day, 0.5 L TB media supplemented with ampicillin, was inoculated with the starter culture and incubated at 37 °C, 250 rpm until the OD_{600} reached 1.0. Cultures were equilibrated at 20 °C for 1 h. Expression was induced by addition of IPTG (final concentration 200 μM). Cultures were incubated at 20 °C, 200 rpm for 18 h.

Overexpression of AnaA and TycF: plasmids containing the gene of interest were transformed into chemically competent BL21(DE3) (*holo*-AnaD: BAP1) *E. coli* cells. A single colony was picked to inoculate a 5 mL LB starter culture grown overnight at 37 °C, 200 rpm. The following day, 1 L LB media supplemented with either kanamycin or ampicillin, was inoculated with the starter culture and incubated at 37 °C, 250 rpm until the OD_{600} reached 0.6. For *apo*-ACP, the media was also supplemented with a trace metals mixture (final concentrations: 50 μM $FeCl_3$, 20 μM $CaCl_2$, 10 μM $MnCl_2$, 10 μM $ZnCl_2$, 2 μM $CoCl_2$, 2 μM $CuCl_2$, 2 μM $NiCl_2$, 2 μM Na_2MoO_4, 2 μM Na_2SeO_3, 2 μM $B(OH)_3$). Cultures were cooled to rt. Expression was induced by addition of IPTG (final concentration 200 μM). Cultures were incubated at 20 °C, 200 rpm for 18 h. AnaB was grown at 25 °C for 18 h instead of 20 °C.

Purification of CoaADE: Cell pellets were resuspended in 4 mL of lysis buffer (100 mM Tris-HCl, 500 mM NaCl, and 0.1 mM DTT at pH 7.9) per gram of wet cell mass and incubated with lysozyme and DNase (Promega) on ice for 1 h. Cells were lysed by sonication (3 s on, 6 s off, 5

min total). Insoluble material was removed by centrifugation (17,000 x *g* for 30 min at 4 °C). The clarified lysate was incubated with gentle shaking along with 5 mL of amylose resin for 1 h at 4 °C and poured into a column. The resin-bound protein was washed with 25 mL lysis buffer and eluted with 16 mL elution buffer (10 mM Tris-HCl, 500 mM NaCl, 0.1 mM DTT, 10 mM maltose, pH 7.9). The eluted protein was concentrated using 3 – 50 kDa centrifugal cutoff filters and dialyzed overnight into lysis buffer. The final protein was concentrated further using 50 kDa centrifugal cutoff filters, aliquoted, flash frozen in liquid nitrogen and stored at -80 °C.

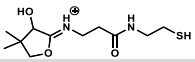
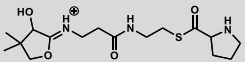
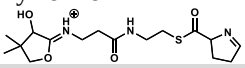
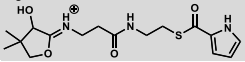
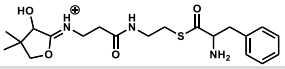
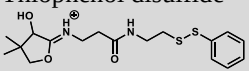
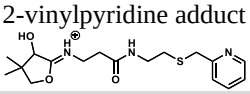
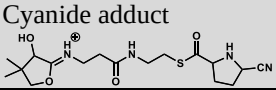
Purification of AnaA and TycF: Cell pellets were resuspended in 4 mL of lysis buffer (50 mM Tris-HCl, 300 mM NaCl, 10 mM imidazole, and 10% (v/v) glycerol at pH 7.4) per gram of wet cell mass. Cells were lysed by sonication (3 s on, 6 s off, 3 min total). Insoluble material was removed by centrifugation (30,000 x *g* for 30 min at 4 °C). The clarified lysate was incubated with gentle shaking along with 1 mL of nickel-NTA resin for 1 h at 4 °C and poured into a column. The resin-bound protein was washed with 25 mL wash buffer (50 mM Tris-HCl, 300 mM NaCl, 25 mM imidazole, 10% (v/v) glycerol, pH 7.4). The desired protein was eluted with 4 mL elution buffer ((50 mM Tris-HCl, 300 mM NaCl, 300 mM imidazole, and 10% (v/v) glycerol at pH 7.4The eluted protein was concentrated using 30 kDa centrifugal cutoff filters and exchanged into storage buffer (50 mM Tris-HCl and 10% (v/v) glycerol at pH 7.4) using a PD-10 column. The desalted protein was concentrated further using 30 kDa centrifugal cutoff filters, aliquoted, flash frozen in liquid nitrogen and stored at -80 °C.

III. Enzymatic reactions and data

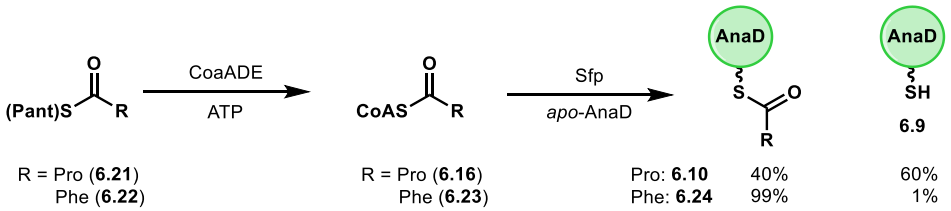
A. General information

Electrospray liquid chromatography-mass spectrometry (LC-MS) analysis was performed on an Agilent G6545A quadrupole-time of flight mass spectrometer in positive mode with an Agilent 1290 UPLC system. Solvent A = water with 0.1% formic acid. Solvent B = 95% acetonitrile, 5% water and 0.1% formic acid. MS data were analyzed using the Agilent Qualitative Mass Hunter software. Intact proteins were deconvoluted with the maximum entropy algorithm.

Recombinant *holo*-AcpP from *E. coli* was prepared by Dr. Gregory Dodge.

Species	Calculated protein mass (Da)[a]	Observed mass (Da)	Calculated Ppant ejection fragment (*m/z*)	Observed Ppant ejection mass (*m/z*)[b]
Apo-AnaD	13,809.8	13,810.2	N/A	N/A
Holo-AnaD 	14,149.9	14,150.4	261.13	261.13
Pro-AnaD 	14,247	14,247	358.18	358.18
Pyrroline 	14,245	14,245	356.16	356.15
Pyrrole 	14,243	14,243	354.15	354.15
Phe-AnaD 	14,297	14,297	408.20	408.20
Thiophenol disulfide 	14,258	14,258	369.13	-
2-vinylpyridine adduct 	14,241	14,241	352.17	352.17
Cyanide adduct 	14,272	ND	383.18	383.18

[a]All constructs from pET28 were observed to lose their initial Met. This also allowed for significant formation of gluconoylated adducts[44] (+178 Da when deconvoluted) that did not appear to affect protein function and were factored into all calculations of PCP species breakdowns. [b]The Ppant ejection assay was not run for all experiments.

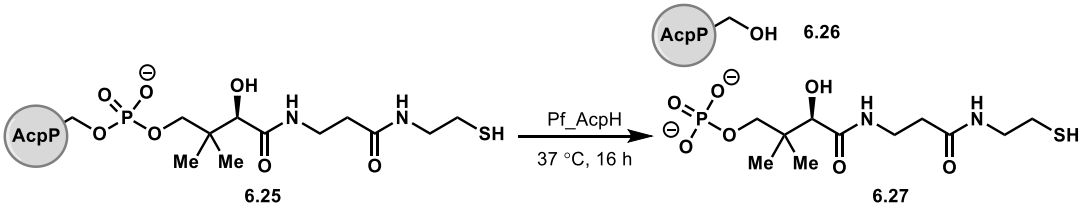


Enzymatic loading of *apo*-AnaD (Figure 6.4A)

Directly from aminoacyl-CoA: 50 µM *apo*-AnaD, 200 µM Pro- or Phe-pantetheine, 10 µM Sfp, in 50 mM Tris-HCl pH 7.5 and 10 mM $MgCl_2$ were combined (total volume 50 µL) and incubated at 37 °C for 1 h.

With CoA biosynthesis: 50 µM *apo*-AnaD, 200 µM Pro- or Phe-CoA, 800 µM ATP, 0.2 mg/mL CoaA, 0.2 mg/mL CoaD, 0.2 mg/mL CoaE, 10 µM Sfp in 50 mM Tris-HCl pH 7.5 and 10 mM $MgCl_2$ were combined (total volume 50 µL) and incubated at 37 °C for 1 h.

In both cases, reactions were quenched by diluting 10x with 1% v/v of formic acid in water, and pelleting any precipitation by centrifuging at 10,000 x *g*, 4 °C for 10 min. The supernatant was transferred to sample vials for LC-MS analysis: column = Phenomenex Aeris 3.6 µm WIDEPORE C4 2.1 x 50 mm; method = 5% B at 0.5 mL/min for 2 min, followed by a linear gradient to 100% B over 4 min, 100% B for 2 min, followed by a 0.1 min linear gradient to 5% B and 1.9 min equilibration at 5% B (total time 10 min). t_R = 4.4 min. The relative abundance of the deconvoluted protein species were used to calculate the fraction of *holo*-, Pro- and Phe-AnaD present.



Conversion of *holo*-CPs to *apo* (Figure 6.4B): 50 µM *holo*-AnaD or AcpP and 2 µM Pf_AcpH in 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 12.5 mM $MgCl_2$, 1 mM $MnCl_2$, and 10% v/v glycerol were combined (total volume 50 µL) and incubated at 37 °C for 16 h. Reactions were quenched by diluting 10x with 1% v/v of formic acid in water, and pelleting any precipitation by centrifuging at 10,000 x *g*, 4 °C for 10 min. The supernatant was transferred to sample vials for LC-MS analysis: column = Phenomenex Aeris 3.6 µm WIDEPORE C4 2.1 x 50 mm; method = 5% B at 0.5 mL/min for 2 min, followed by a linear gradient to 100% B over 4 min, 100% B for 2 min, followed by a 0.1 min linear gradient to 5% B and 1.9 min equilibration at 5% B (total time 10 min). t_R (AcpP) = 4.1 min. The relative abundance of the deconvoluted protein species were used to calculate the fraction of *holo*- and *apo*-CP.

Holo-AcpP was completely converted, but we observed no activity on *holo*-AnaD.

Figure 6.S2. Sequence alignment of AnaD and AcpP.
Sequences were aligned by MUSCLE (https://www.ebi.ac.uk/Tools/msa/muscle/) and colored in JalView.

Diffusive loading of holo-AnaD (Figure 6.5): 50 µM *holo*-AnaD, 1 mM aminoacyl-SNAC or -SPh, in 50 mM potassium phosphate pH 6.8 were combined (total volume 50 µL) and incubated at 28 °C for 1-3 h. The total loading was calculated by comparing the area of deconvoluted protein peaks after intact protein MS. 2-vinylpyridine (8 mM) formed unproductive adducts with *holo*-AnaD (61% of all CP present after 4 h) and was omitted from later reactions.

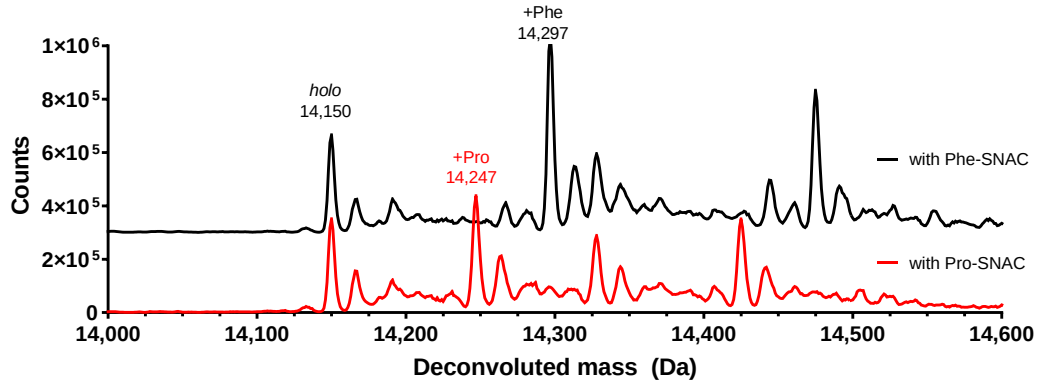
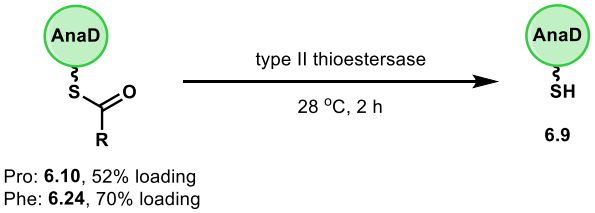


Figure 6.S3. Deconvoluted protein spectra of *holo*-AnaD incubated with aminoacyl-SNAC thioesters.



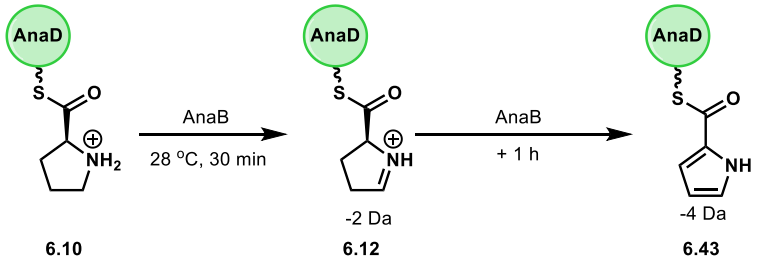
Hydrolysis of aminoacyl groups from AnaD (Table 6.1): 50 µM AnaD (mixture of *holo* and aminoacyl) and 2 µM thioesterase in 50 mM potassium phosphate pH 6.8 were combined (total volume 50 µL) and incubated at 28 °C for 1-3 h. The total hydrolysis was calculated by comparing

the area of deconvoluted protein peaks after intact protein MS. HSNAC-accelerated hydrolysis was measured by adding 1 mM HSNAC to reactions.

Hydrolysis of aminoacyl-SNACs: 1 mM Pro- or Phe-SNAC and 2 µM TycF in 50 mM potassium phosphate pH 6.8 were combined (total volume 100 µL) and incubated at 28 °C for 4 h. After every hour, 25 µL was removed and quenched in 75 µL of MeOH. Precipitate was pelleted by centrifuging at 10,000 x *g* for 10 min. The supernatant was transferred to sample vials for LC-MS analysis: column = Waters Acquity 1.7 µm UPLC BEH C18 column, 2.1 mm x 50 mm; method = 5% B at 0.5 mL/min for 30 s, followed by a linear gradient to 100% B over 3 min, 100% B for 1 min, followed by a 0.1 min linear gradient to 5% B and 0.4 min equilibration at 5% B (total time 5 min). Minor hydrolysis was observed in the no-enzyme negative controls.

Pro-SNAC: 6% total hydrolysis after 4 h

Phe-SNAC: 20% total hydrolysis after 4 h



6.10 6.12 6.43

AnaB-catalyzed oxidation of Pro-AnaD (Figure 6.7): 50 µM Pro-AnaD (mixture of *holo* and Pro) and 2 µM AnaB in 50 mM potassium phosphate pH 6.8 were combined (total volume 75 µL) and incubated at 28 °C for 1.5 h. Every 30 min, 25 µL aliquots were removed and quenched by diluting 10x in 1% v/v formic acid in water. The oxidation was calculated by comparing the intensity of the ejected Ppant ions. Because of the small mass changes, the deconvoluted protein

peak appeared to be a mixture of the three species, with the molecular weight becoming closer to the pyrrole mass over time.

On free small molecules: other substrates tested included L-Pro, Pro-SNAC, Pro-pantetheine, Pro-phosphopantetheine (via CoaA) and Pro-CoA (via CoaADE). No oxidation was observed by small molecule MS. For pantetheine and larger molecules: column = Waters XBridge C18 3.5 µm , 2.1 x 150 mm; solvent A = water with 0.1% v/v formic acid, solvent B = methanol with 0.1% v/v formic acid; method = 10% B at 0.2 mL/min for 2 min, followed by a linear gradient to 90% B over 15 min, then a gradient to 90% over 1 min (total time 18 min). t_R (Pro-pantetheine) = 8.2 min; t_R (Pro-phosphopantetheine) = 5.2 min; t_R (3'-dephospho-Pro-CoA) = 5.9 min.

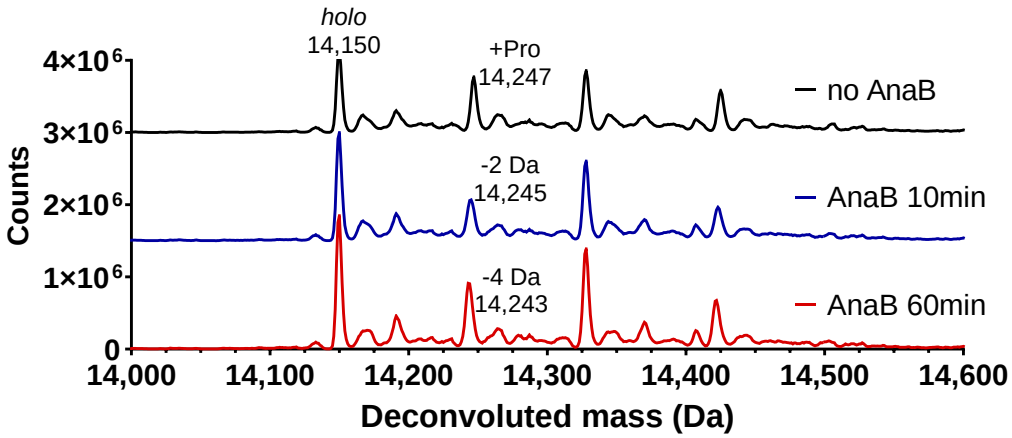
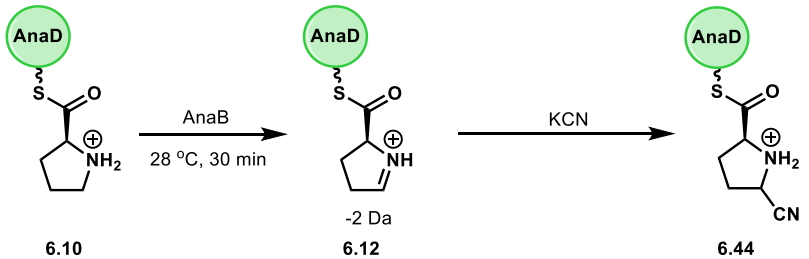


Figure 6.S4. Deconvoluted protein spectra of Pro-AnaD incubated with AnaB.



AnaB-catalyzed oxidation with nucleophilic addition: 50 µM Pro-AnaD (mixture of *holo* and Pro), 1 mM potassium cyanide and 1.5 µM AnaB in 50 mM potassium phosphate pH 6.8 were combined (total volume 50 µL) and incubated at 28 °C for 2 h. Every hour, 25 µL aliquots were

removed and quenched by diluting 10x in 1% v/v formic acid in water. The oxidation and nucleophilic addition were observed by comparing the intensity of the ejected Ppant ions.

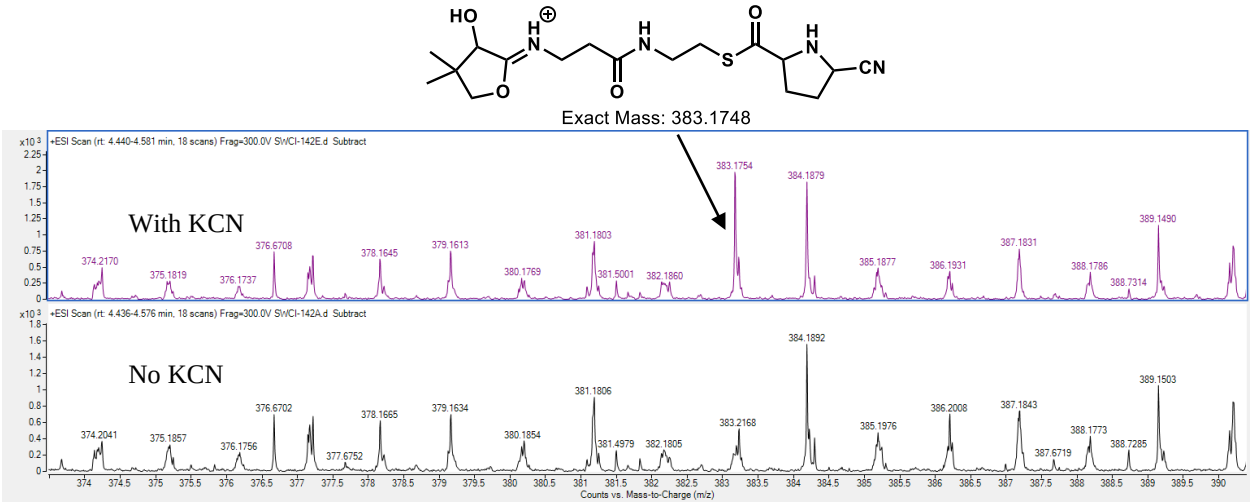
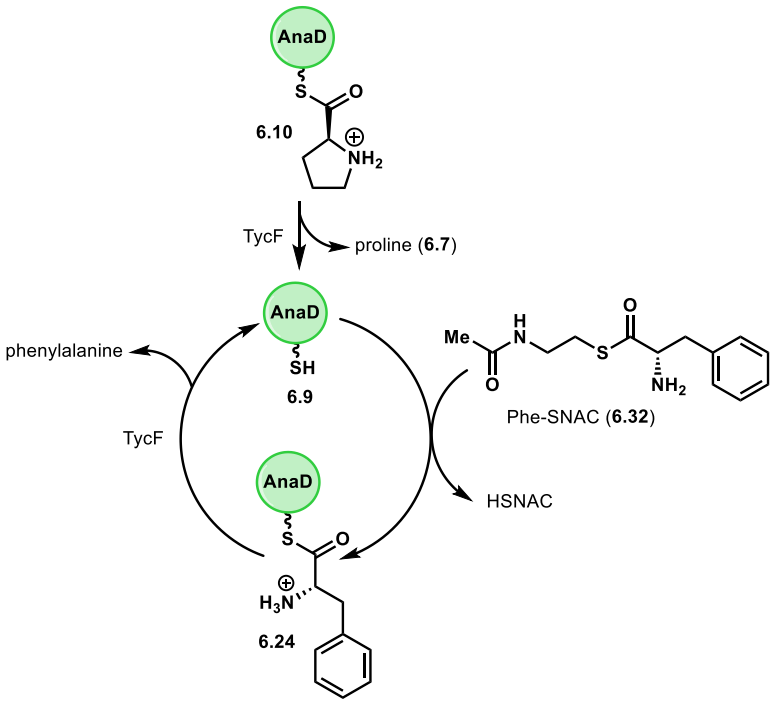
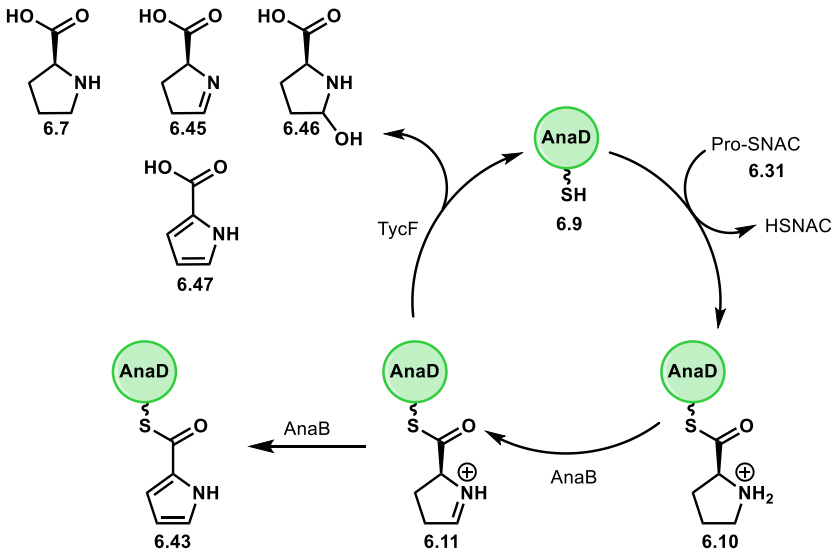


Figure 6.S5. Ejected ions of the AnaB oxidation reaction with and without KCN present. A new peak is present in the KCN reaction matching the mass of a cyanide addition ejected ion. While this peak did not grow over time, it also did not appear to be oxidized a second time by AnaB.



One-pot offloading and loading (Table 6.2): 50 µM Pro-AnaD (mixture of *holo* and Pro), 1 mM Phe-SNAC and 2 µM TycF in 50 mM potassium phosphate pH 6.8 were combined (total volume

100 µL) and incubated at 28 °C for 1.5 h. Every hour, 25 µL aliquots were removed and quenched by diluting 10x in 1% v/v formic acid in water. The distribution of aminoacyl and *holo* species was calculated by comparing the area of deconvoluted protein peaks after intact protein MS.



One-pot loading, oxidation and offloading (Figure 6.8): 50 µM *holo*-AnaD, 1 mM Pro-SNAC, 1.5 µM AnaB, 2 µM TycF and a 50 µM phenylalanine internal standard in 50 mM potassium phosphate pH 6.8 were combined (total volume 200 µL) and incubated at 28 °C for 4 h. Every hour, 50 µL aliquots were removed and quenched by diluting 10x in 1% v/v formic acid in water. Precipitation was pelleted by centrifuging 10,000 x *g* for 10 min. Half of the supernantant was transferred to sample vials for intact protein MS. The distribution of aminoacyl and *holo* species was calculated by comparing the area of deconvoluted protein peaks after intact protein MS. The other half was filtered through a 3 kDa molecular cutoff filter (Millipore) and the flow-through analyzed by small molecule LC-MS: column = Waters Acquity 1.7 µm UPLC BEH C18 column, 2.1 mm x 50 mm; method = 5% B at 0.5 mL/min for 30 s, followed by a linear gradient to 100% B over 3 min, 100% B for 1 min, followed by a 0.1 min linear gradient to 5% B and 0.4 min

equilibration at 5% B (total time 5 min). t_R (**6.45**) = 0.9 min; t_R (phenylalanine) = 1.2 min. The area of the extracted ion chromatograms of acid **6.45** vs. phenylalanine were compared.[45]

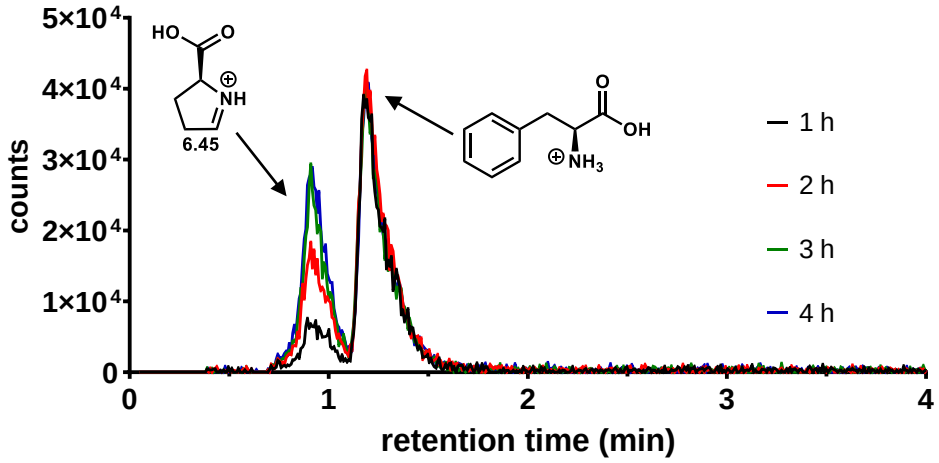
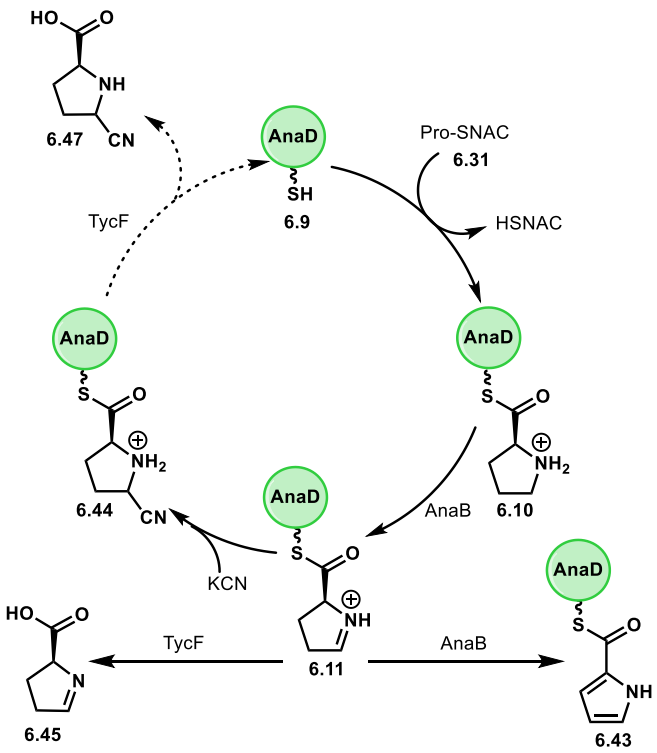


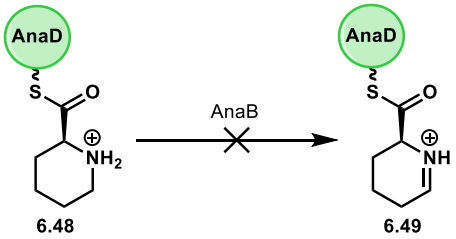
Figure 6.S6. Extracted ion chromatograms of the pyrrolinium acid product relative to the internal standard.

Pro-AnaD and AnaB were also incubated for 2 h for complete conversion to the pyrrole **6.44** before addition of TycF. There was no cleavage to regenerate *holo*-AnaD. No transthioesterification occurred when *holo*-AnaD was incubated with thioester donor **6.S9**.

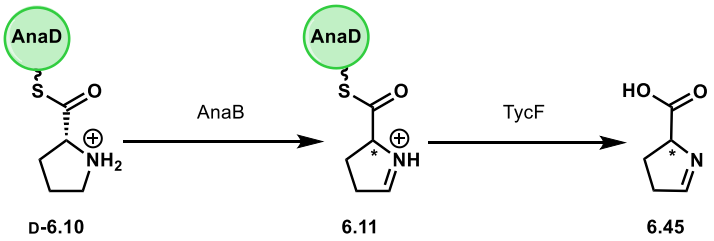


One-pot loading, oxidation, nucleophilic addition and offloading (Figure 6.9): 2 mM KCN, 50 µM *holo*-AnaD, 1 mM Pro-SNAC, 1.5 µM AnaB, 2 µM TycF and a 50 µM phenylalanine internal standard in 50 mM potassium phosphate pH 6.8 were combined (total volume 200 µL) and incubated at 28 °C for 4 h. Every hour, 50 µL aliquots were removed and quenched by diluting 10x in 1% v/v formic acid in water. Half of the quenched sample was analyzed by intact protein MS, and the other half by small molecule MS after filtering out proteins.

We detected oxidation to **6.43** and a small quantity of acid **6.45**, but no aminoacyl-AnaD **6.44** or acid **6.47**.



Oxidation of a six-membered ring (Figure 6.10): 50 µM *apo*-AnaD, 200 µM thioester **6.S10**, 800 µM ATP, 0.2 mg/mL CoaA, 0.2 mg/mL CoaD, 0.2 mg/mL CoaE, 10 µM Sfp and 2 µM AnaB, in 50 mM potassium phosphate pH 6.8 and 10 mM $MgCl_2$ were combined (total volume 50 µL) and incubated at 28 °C for 3 h. The reactions were analyzed by intact protein MS. We observed 59% loading, but no oxidation, both by deconvoluted protein peaks and ejected Ppant ions.



Oxidation of D-Pro-AnaD (Figure 6.10): 50 µM *holo*-AnaD, 2 mM D-Pro-SNAC, 5 µM AnaB, 2 µM TycF and a 50 µM phenylalanine standard in 50 mM potassium phosphate pH 6.8 were combined (total volume 100 µL) and incubated at 28 °C for 3 h. The reactions were analyzed by intact protein MS to detect PCP-bound intermediates and small molecule MS for acid products. Oxidation by AnaB was slower, and low quantities (seeming less than the L-Pro reaction, but not rigorously quantified) of acid **6.45** were observed.

6.6 References

(1) James, K. J.; Crowley, J.; Duphard, J.; Lehane, M.; Furey, A. Anatoxin-a and Analogues: Discovery, Distribution, and Toxicology. In *Phycotoxins: Chemistry and Biochemistry*; Blackwell Publishing: Ames, Iowa, USA, 2007; pp 141–158.

(2) Devlin, J. P.; Edwards, O. E.; Gorham, P. R.; Hunter, N. R.; Pike, R. K.; Stavric, B. Anatoxin-a, a Toxic Alkaloid from Anabaena Flos-Aquae NRC-44h. *Can. J. Chem.* **1977**, *55*, 1367–1371.

(3) Aronstam, R. S.; Witkop, B. Anatoxin-a Interactions with Cholinergic Synaptic Molecules. *Proc. Natl. Acad. Sci.* **1981**, *78*, 4639–4643.

(4) Fiore, M. F.; de Lima, S. T.; Carmichael, W. W.; McKinnie, S. M. K.; Chekan, J. R.; Moore, B. S. Guanitoxin, Re-Naming a Cyanobacterial Organophosphate Toxin. *Harmful Algae* **2020**, *92*, 101737.

(5) Stevens, D. K.; Krieger, R. I. Stability Studies on the Cyanobacterial Nicotinic Alkaloid Snatoxin-A. *Toxicon* **1991**, *29*, 167–179.

(6) Munday, R.; Thomas, K.; Gibbs, R.; Murphy, C.; Quilliam, M. A. Acute Toxicities of Saxitoxin, Neosaxitoxin, Decarbamoyl Saxitoxin and Gonyautoxins 1&4 and 2&3 to Mice by Various Routes of Administration. *Toxicon* **2013**, *76*, 77–83.

(7) Hemscheidt, T.; Rapala, J.; Sivonen, K.; Skulberg, O. M. Biosynthesis of Anatoxin-a in Anabaena Flos-Aquae and Homoanatoxin-a in Oscillatoria Formosa. *J. Chem. Soc. Chem. Commun.* **1995**, No. 13, 1361.

(8) Cadel-Six, S.; Iteman, I.; Peyraud-Thomas, C.; Mann, S.; Ploux, O.; Méjean, A. Identification of a Polyketide Synthase Coding Sequence Specific for Anatoxin-a-Producing Oscillatoria Cyanobacteria. *Appl. Environ. Microbiol.* **2009**, *75*, 4909–4912.

(9) Méjean, A.; Mann, S.; Maldiney, T.; Vassiliadis, G.; Lequin, O.; Ploux, O. Evidence That Biosynthesis of the Neurotoxic Alkaloids Anatoxin-a and Homoanatoxin-a in the Cyanobacterium Oscillatoria PCC 6506 Occurs on a Modular Polyketide Synthase Initiated by l -Proline. *J. Am. Chem. Soc.* **2009**, *131*, 7512–7513.

(10) Méjean, A.; Mann, S.; Vassiliadis, G.; Lombard, B.; Loew, D.; Ploux, O. In Vitro Reconstitution of the First Steps of Anatoxin-a Biosynthesis in Oscillatoria PCC 6506: From Free L-Proline to Acyl Carrier Protein Bound Dehydroproline. *Biochemistry* **2010**, *49*, 103–113.

(11) Mann, S.; Lombard, B.; Loew, D.; Méjean, A.; Ploux, O. Insights into the Reaction Mechanism of the Prolyl - Acyl Carrier Protein Oxidase Involved in Anatoxin-a and Homoanatoxin-a Biosynthesis. *Biochemistry* **2011**, *50*, 7184–7197.

(12) Roujeinikova, A.; Baldock, C.; Simon, W. J.; Gilroy, J.; Baker, P. J.; Stuitje, A. R.; Rice, D. W.; Slabas, A. R.; Rafferty, J. B. X-Ray Crystallographic Studies on Butyryl-ACP Reveal Flexibility of the Structure around a Putative Acyl Chain Binding Site. *Structure* **2002**, *10*, 825–835.

(13) Ghislieri, D.; Turner, N. J. Biocatalytic Approaches to the Synthesis of Enantiomerically Pure Chiral Amines. *Top. Catal.* **2014**, *57*, 284–300.

(14) Nugent, T. C.; El-Shazly, M. Chiral Amine Synthesis - Recent Developments and Trends for Enamide Reduction, Reductive Amination, and Imine Reduction. *Advanced Synthesis and Catalysis*. March 22, 2010, pp 753–819.

(15) Worthington, A. S.; Burkart, M. D. One-Pot Chemo-Enzymatic Synthesis of Reporter-Modified Proteins. *Org. Biomol. Chem.* **2006**, *4*, 44–46.

(16) Hansen, D. A.; Koch, A. A.; Sherman, D. H. Substrate Controlled Divergence in Polyketide

Synthase Catalysis. *J. Am. Chem. Soc.* **2015**, *137*, 3735–3738.
(17) Quadri, L. E. N.; Weinreb, P. H.; Lei, M.; Nakano, M. M.; Zuber, P.; Walsh, C. T. Characterization of Sfp, a Bacillus Subtilis Phosphopantetheinyl Transferase for Peptidyl Carrier Protein Domains in Peptide Synthetases. *Biochemistry* **1998**, *37*, 1585–1595.
(18) Kosa, N. M.; Haushalter, R. W.; Smith, A. R.; Burkart, M. D. Reversible Labeling of Native and Fusion-Protein Motifs. *Nat. Methods* **2012**, *9*, 981–984.
(19) Pfeifer, B. A.; Admiraal, S. J.; Gramajo, H.; Cane, D. E.; Khosla, C. Biosynthesis of Complex Polyketides in a Metabolically Engineered Stain of E. Coli. *Science* **2001**, *291*, 1790–1792.
(20) Pates, G. O.; Guler, L.; Nash, J. J.; Kenttämaa, H. I. Reactivity and Selectivity of Charged Phenyl Radicals toward Amino Acids in a Fourier Transform Ion Cyclotron Resonance Mass Spectrometer. *J. Am. Chem. Soc.* **2011**, *133*, 9331–9342.
(21) Thomas, J.; Cronan, J. E. The Enigmatic Acyl Carrier Protein Phosphodiesterase of Escherichia Coli. *J. Biol. Chem.* **2005**, *280*, 34675–34683.
(22) Murugan, E.; Kong, R.; Sun, H.; Rao, F.; Liang, Z. X. Expression, Purification and Characterization of the Acyl Carrier Protein Phosphodiesterase from Pseudomonas Aeruginosa. *Protein Expr. Purif.* **2010**, *71*, 132–138.
(23) Kosa, N. M.; Pham, K. M.; Burkart, M. D. Chemoenzymatic Exchange of Phosphopantetheine on Protein and Peptide. *Chem. Sci.* **2014**, *5*, 1179–1186.
(24) Amann, E.; Ochs, B.; Abel, K.-J. Tightly Regulated Tac Promoter Vectors Useful for the Expression of Unfused and Fused Proteins in Escherichia Coli. *Gene* **1988**, *69*, 301–315.
(25) Kotowska, M.; Pawlik, K. Roles of Type II Thioesterases and Their Application for Secondary Metabolite Yield Improvement. *Appl. Microbiol. Biotechnol.* **2014**, *98*, 7735–7746.
(26) Yeh, E.; Kohli, R. M.; Bruner, S. D.; Walsh, C. T. Type II Thioesterase Restores Activity of a NRPS Module Stalled with an Aminoacyl-S-Enzyme That Cannot Be Elongated. *ChemBioChem* **2004**, *5*, 1290–1293.
(27) Whicher, J. R.; Florova, G.; Sydor, P. K.; Singh, R.; Alhamadsheh, M.; Challis, G. L.; Reynolds, K. A.; Smith, J. L. Structure and Function of the RedJ Protein, a Thioesterase from the Prodiginine Biosynthetic Pathway in Streptomyces Coelicolor. *J. Biol. Chem.* **2011**, *286*, 22558–22569.
(28) Claxton, H. B.; Akey, D. L.; Silver, M. K.; Admiraal, S. J.; Smith, J. L. Structure and Functional Analysis of RifR, the Type II Thioesterase from the Rifamycin Biosynthetic Pathway. *J. Biol. Chem.* **2009**, *284*, 5021–5029.
(29) Koglin, A.; Löhr, F.; Bernhard, F.; Rogov, V. V.; Frueh, D. P.; Strieter, E. R.; Mofid, M. R.; Güntert, P.; Wagner, G.; Walsh, C. T.; Marahiel, M. A.; Dötsch, V. Structural Basis for the Selectivity of the External Thioesterase of the Surfactin Synthetase. *Nature* **2008**, *454*, 907–911.
(30) Nowak-Thompson, B.; Chaney, N.; Wing, J. S.; Gould, S. J.; Loper, J. E. Characterization of the Pyoluteorin Biosynthetic Gene Cluster of Pseudomonas Fluorescens Pf-5. *J. Bacteriol.* **1999**, *181*, 2166–2174.
(31) Berditsch, M.; Trapp, M.; Afonin, S.; Weber, C.; Misiewicz, J.; Turkson, J.; Ulrich, A. S. Antimicrobial Peptide Gramicidin S Is Accumulated in Granules of Producer Cells for Storage of Bacterial Phosphagens. *Sci. Rep.* **2017**, *7*, 1–11.
(32) Thomas, M. G.; Burkart, M. D.; Walsh, C. T. Conversion of L-Proline to Pyrrolyl-2-Carboxyl-S-PCP during Undecylprodigiosin and Pyoluteorin Biosynthesis. *Chem. Biol.*

2002, *9*, 171–184.
(33) Lee, C.-C.; Ko, T.-P.; Chen, C.-T.; Chan, Y.-T.; Lo, S.-Y.; Chang, J.-Y.; Chen, Y.-W.; Chung, T.-F.; Hsieh, H.-J.; Hsiao, C.-D.; Wang, A. H. J. Crystal Structure of PigA: A Prolyl Thioester-Oxidizing Enzyme in Prodigiosin Biosynthesis. *ChemBioChem* **2019**, *20*, 193–202.
(34) Garneau, S.; Dorrestein, P. C.; Kelleher, N. L.; Walsh, C. T. Characterization of the Formation of the Pyrrole Moiety during Clorobiocin and Coumermycin A1 Biosynthesis. *Biochemistry* **2005**, *44*, 2770–2780.
(35) Méjean, A.; Paci, G.; Gautier, V.; Ploux, O. Biosynthesis of Anatoxin-a and Analogues (Anatoxins) in Cyanobacteria. *Toxicon* **2014**, *91*, 15–22.
(36) Meluzzi, D.; Zheng, W. H.; Hensler, M.; Nizet, V.; Dorrestein, P. C. Top-down Mass Spectrometry on Low-Resolution Instruments: Characterization of Phosphopantetheinylated Carrier Domains in Polyketide and Non-Ribosomal Biosynthetic Pathways. *Bioorganic Med. Chem. Lett.* **2008**, *18*, 3107–3111.
(37) Moncoq, K.; Regad, L.; Mann, S.; Méjean, A.; Ploux, O. Structure of the Prolyl-Acyl Carrier Protein Oxidase Involved in the Biosynthesis of the Cyanotoxin Anatoxin-A. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2013**, *69*, 2340–2352.
(38) Katritzky, A.; Shestopalov, A.; Suzuki, K. A New Convenient Preparation of Thiol Esters Utilizing N -Acylbenzotriazoles. *Synthesis (Stuttg).* **2004**, *2004*, 1806–1813.
(39) Gay, D.; You, Y. O.; Keatinge-Clay, A. T.; Cane, D. E. Structure and Stereospecificity of the Dehydratase Domain from the Terminal Module of the Rifamycin Polyketide Synthase. *Biochemistry* **2013**, *52*, 8916–8928.
(40) Luo, N.; Yang, Y.-B.; Yang, X.-Q.; Miao, C.-P.; Li, Y.-Q.; Xu, L.-H.; Ding, Z.-T.; Zhao, L.-X. The Streptazolin- and Obscurolide-Type Metabolites from Soil-Derived Streptomyces Alboniger YIM20533 and the Mechanism of Influence of γ-Butyrolactone on the Growth of Streptomyces by Their Non-Enzymatic Reaction Biosynthesis. *RSC Adv.* **2018**, *8*, 35042–35049.
(41) Hansen, D. A.; Rath, C. M.; Eisman, E. B.; Narayan, A. R. H.; Kittendorf, J. D.; Mortison, J. D.; Yoon, Y. J.; Sherman, D. H. Biocatalytic Synthesis of Pikromycin, Methymycin, Neomethymycin, Novamethymycin, and Ketomethymycin. *J. Am. Chem. Soc.* **2013**, *135*, 11232–11238.
(42) Naveen, S.; Dinesh, B.; Abiraj, K.; Channe Gowda, D.; Sridhar, M. A.; Shashidhara Prasad, J. Synthesis and Crystal Structure of Tert-Butyl 2-((Phenylthio)-Carbonyl) Pyrrolidine-1-Carboxylate. *J. Chem. Crystallogr.* **2007**, *37*, 721–725.
(43) Eschenfeldt, W. H.; Lucy, S.; Millard, C. S.; Joachimiak, A.; Mark, I. D. A Family of LIC Vectors for High-Throughput Cloning and Purification of Proteins. In *Methods in Molecular Biology*; Doyle, S. A., Ed.; Humana Press: Totowa, NJ, 2009; Vol. 498, pp 105–115.
(44) Geoghegan, K. F.; Dixon, H. B. F.; Rosner, P. J.; Hoth, L. R.; Lanzetti, A. J.; Borzilleri, K. A.; Marr, E. S.; Pezzullo, L. H.; Martin, L. B.; Lemotte, P. K.; Mccoll, A. S.; Kamath, A. V.; Stroh, J. G. Spontaneous Alpha-N-6-Phosphogluconoylation of a "His Tag" in Escherichia Coli: The Cause of Extra Mass of 258 or 178 Da in Fusion Proteins. *Anal. Biochem.* **1999**, *267*, 169–184.
(45) Martinez, S.; Hausinger, R. P. Biochemical and Spectroscopic Characterization of the Non-Heme Fe(II)- and 2-Oxoglutarate-Dependent Ethylene-Forming Enzyme from Pseudomonas Syringae Pv. Phaseolicola PK2. *Biochemistry* **2016**, *55*, 5989–5999.

Chapter 7: Conclusions and Future Directions

Enzymes from carrier protein-dependent biosynthetic pathways catalyze a wealth of interesting transformations but have untapped synthetic potential in the formation of new molecules. Because of the CP-dependence, biocatalytic platforms for PEs will require removing the dependence, exploiting CP-independent reaction modes or adding loading/offloading steps to recycle CPs. This Chapter summarizes the results of the five previous Chapters detailing studies on two PEs, SxtA and AnaB, to generate chiral amines chemoenzymatically.

SxtA is a four-domain polyketide-like synthase which had previously been proposed to initiate biosynthesis of the potent paralytic shellfish toxin saxitoxin.[1] Bioinformatic searches indicated that the SxtA module possesses an α-oxoamine synthase domain, which could serve as a potential tool to modify amino acids. We identified the true starter unit of SxtA and characterized its native activity, facilitating further synthetic and mechanistic studies with individual domains. We have found that SxtA AONS is a promiscuous enzyme for the generation of α-amino ketones in a single step, but with overall robustness that requires improvement by directed evolution. The domain performs many more protonation/deprotonation turnovers, enabling high levels of deuterium incorporation into α-amino acids and their methyl esters and a chemoenzymatic synthesis of an isotopically labeled drug.

SxtA AONS binds and deprotonates the α-carbon of all linear α-amino methyl esters screened, but not the cyclic amino acid, L-proline. AnaB, a previously characterized PE-dependent oxidase,[2] provides an opportunity to derivatize this amino acid and efficiently generate chiral cyclic amines

through an artificial four-step CP-PE catalytic cycle. The full one-pot cycle cannot be completed efficiently due to overoxidation to an unreactive byproduct. These initial studies to explore PE-mediated non-native reactions both in the presence or absence of CPs will contribute to the development of new biocatalytic tools for the synthesis of complex molecules.

Future Directions

Engineering Dimethylation Activity in SxtA MT

We were not able identify the residues in SxtA MT that govern the monomethylation activity compared to the dimethylation observed in other metal-dependent methyltransferases identified in PKS initiation modules (MT_Ls).[3] Additional variants of SxtA MT-GNAT outside the first sphere of interactions with malonyl-ACP will be generated to screen for dimethylating variants. These results will contribute to the growing body of literature on MT_L domains and natural products containing ethyl, isopropyl or *t*-butyl substituents. For example, the bioactivity of molecules at one level of methylation (e.g., one group added in saxitoxin, myxovirescin A)[4] could be tuned with additional methylation or alkylation with SAM derivatives (Figure 7.1). The methyl functionalization activity could also be decreased in the biosynthesis of an originally dimethylated natural product such as gephyronic acid.[5] Methylation could also be introduced into pathways with β-keto-CP intermediates where there is currently no functionalization.

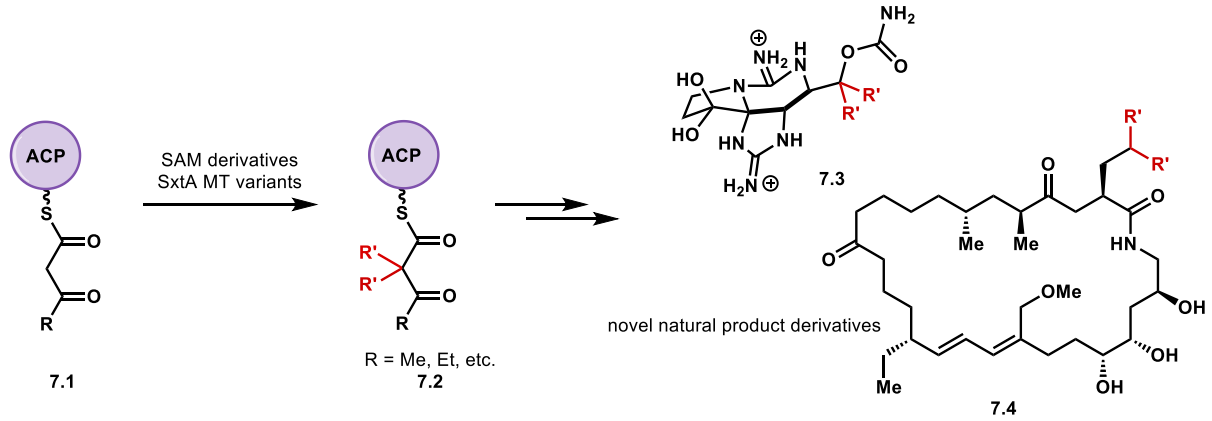


Figure 7.1. Generation of novel natural products through MT_L-catalyzed alkylation.

Generation of α-amino ketones by SxtA AONS

Directed evolution of the excised SxtA AONS domain to engineer increased activity on a non-native Trp substrate is ongoing. These libraries will also be screened for additional activity, such as higher conversion in the native ketone-forming reaction with arginine or other amino acids, or the generation of increased quantities of α-tetra-substituted amino acid derivatives. The small volumes of analytes used in high-throughput droplet screening may allow for miniaturized or droplet enzymatic reactions as well to increase efficiency further. Potential product inhibition may be relieved in cascade reactions removing or elaborating on the α-amino ketone products (e.g. reduction of the carbonyl to a hydroxyl group, see Figure 7.2A).[6]

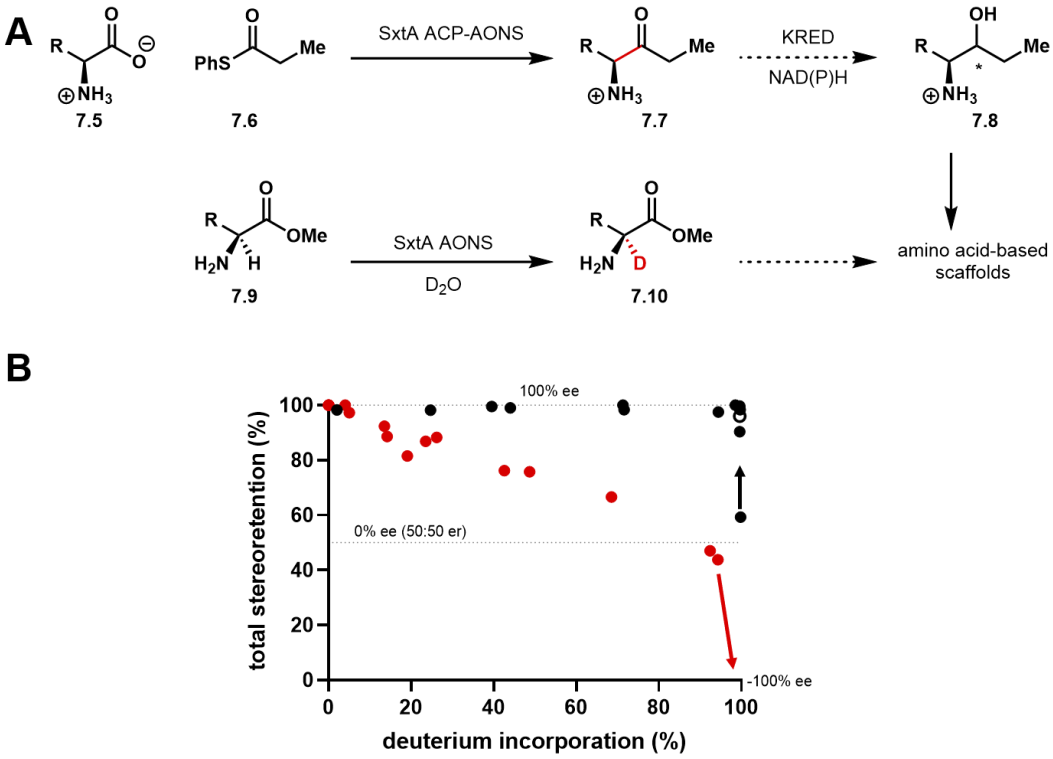


Figure 7.2. Future work on products generated from SxtA AONS.
(A) Synthesis from α-amino ketone or α-deuterated amino acid/methyl ester building blocks. (B) Increasing the stereoselectivity of deuterated methyl esters. Black dots: L-substrates; red: D-substrates.

α-deuteration of amino acids and methyl esters by SxtA AONS

Investigations to understand stereoselectivity of the reprotonation step for each D- and L-methyl ester substrates will continue. Unreactive thioester mimics such as thioethers will be synthesized to increase deuterium incorporation into less active substrates,[7] and we will also screen other α-carbonyl compounds, such as amides. Deuterated products will serve as building blocks in labeled peptides/proteins and other amino acid-based scaffolds. Compounds that are metabolized at the α-carbon may also be screened for changes in half-lives relative to the protio-compounds.

Preliminary experiments have suggested that wild-type SxtA AONS could be a catalyst for stereoinversion or dynamic kinetic resolution of some α-amino methyl esters (Figure 7.2B). Wider screening of methyl esters will identify substrates for this application and members of the AONS variant library can be screened for their stereoselectivity.

In addition, we will elucidate the origin of the broader deuterium incorporation of α-amino methyl esters over the corresponding acids by SxtA AONS. Methyl esters possess a decreased pK_a at the α-carbon and it is not clear if the esterification increases deuterium because of the increased acidity or because of structural resemblance to the ethyl ketone product or a combination. Comparison of SxtA AONS deuteration to another AOS, serine palmitoyltransferase, on long chain aliphatic myristyl esters may reveal the overall effect of esterification on binding to the PLP cofactor and deprotonation.

Synthesis of chiral cyclic amines with AnaB and AnaD

The formation of undesired pyrrole byproducts causes the artificial catalytic cycle to stall in multiple ways, halting productive generation of new chiral compounds. Overoxidation may be minimized by screening other AnaB homologs, a strategy that was successful in identifying a parent for directed evolution in SxtA AONS. The desired electrophilic iminium intermediate can

be intercepted more rapidly after screening for more reactive nucleophiles, and then cleaved from AnaD by a promiscuous thioesterase. Cyclic amines generated by a productive four-step catalytic cycle can be elaborated into more complex small molecules.

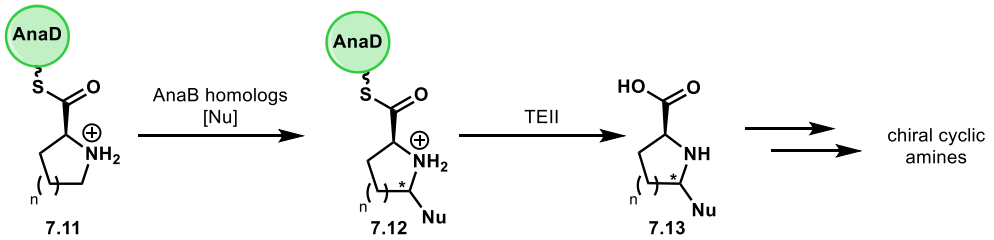


Figure 7.3. Synthesis of chiral cyclic amine building blocks after AnaB-mediated oxidation.

The strategies presented in this thesis, from enzyme and reaction discovery, to chemoenzymatic preparation of valuable compounds demonstrate the possibilities of leveraging CP-dependent PEs in synthesis. We anticipate that these approaches, CP-independent modes of reactivity and CP recycling with inexpensive acyl and aminoacyl donors, will also be applied to other PEs for the efficient, chemo-, site- and stereoselective generation of chiral amines and other important synthetic building blocks.

References

(1) Kellmann, R.; Mihali, T. K.; Young, J. J.; Pickford, R.; Pomati, F.; Neilan, B. A. Biosynthetic Intermediate Analysis and Functional Homology Reveal a Saxitoxin Gene Cluster in Cyanobacteria. *Appl. Environ. Microbiol.* **2008**, *74*, 4044–4053.

(2) Méjean, A.; Mann, S.; Vassiliadis, G.; Lombard, B.; Loew, D.; Ploux, O. In Vitro Reconstitution of the First Steps of Anatoxin-a Biosynthesis in Oscillatoria PCC 6506: From Free L-Proline to Acyl Carrier Protein Bound Dehydroproline. *Biochemistry* **2010**, *49*, 103–113.

(3) Skiba, M. A. Structural and Biochemical Investigation of Methylation and Elucidation of T-Butyl Formation in Polyketide Biosynthesis, University of Michigan, 2018.

(4) Simunovic, V.; Zapp, J.; Rachid, S.; Krug, D.; Meiser, P.; Müller, R. Myxovirescin A Biosynthesis Is Directed by Hybrid Polyketide Synthases/Nonribosomal Peptide Synthetase, 3-Hydroxy-3-Methylglutaryl-CoA Synthases, and Trans-Acting Acyltransferases. *ChemBioChem* **2006**, *7*, 1206–1220.

(5) Young, J.; Stevens, D. C.; Carmichael, R.; Tan, J.; Rachid, S.; Boddy, C. N.; Müller, R.; Taylor, R. E. Elucidation of Gephyronic Acid Biosynthetic Pathway Revealed Unexpected SAM-Dependent Methylations. *J. Nat. Prod.* **2013**, *76*, 2269–2276.

(6) Cao, J.; Hyster, T. K. Pyridoxal-Catalyzed Racemization of α-Aminoketones Enables the Stereodivergent Synthesis of 1,2-Amino Alcohols Using Ketoreductases. *ACS Catal.* **2020**, *10*, 6171–6175.
(7) Ikushiro, H.; Fujii, S.; Shiraiwa, Y.; Hayashi, H. Acceleration of the Substrate Cα Deprotonation by an Analogue of the Second Substrate Palmitoyl-CoA in Serine Palmitoyltransferase. *J. Biol. Chem.* **2008**, *283*, 7542–7553.